

A Parisian museum to be defended

Museums are not the same as galleries, and should not become so. The Musée de l'Homme in Paris presents an excellent opportunity publicly to bring together disciplines that were until recently estranged.

If one agrees with Alexander Pope that the proper study of mankind is man, one should welcome the overhaul of the Musée de l'Homme in Paris as an excellent opportunity to draw together the increasingly disparate strands of research on human beings, and present them to the public, under one roof. Unfortunately, the conflicting temptation (see page 109) to reinvent the museum as a gallery of primitive art also has its attractions (not least financial). To see the scale of the opportunity that would thus be lost, one needs to examine why museums and galleries are different things, and how they have become increasingly so over the past few years.

The essence of a gallery is its provision of space wherein one contemplates objects — usually art — that exist either for their own sake or as representations of aesthetic values more or less remote from scholarship. Many of the objects now housed in the Musée de l'Homme, or in the Paris Museum of African and Oceanic Arts (whose collections are planned to be merged with those of the Musée de l'Homme), are beautiful in themselves. But to display them simply as *objets d'art* would be to miss the chance to exploit their other function, that of the product of a culture and technology the interest of which is greater than the object itself.

There is another danger. The display of artefacts for their own sake is increasingly seen as patronizing. Curators of ethnographic collections face increasing demands from native peoples for the repatriation of objects of ritual significance. Unless one can show that the inclusion of such things in a museum display is in the context of educational purposes, such demands seem justified. In that case, a display of ethnic objects for their own sake is not only outmoded, but is a kind of cross-cultural voyeurism.

Thankfully, museums have moved on from being just galleries — as too many used to be. It is now rare for their displays to consist of obscurely labelled curiosities, exhibited without context. The need to provide the visitor with some idea about specimens' origins and functions is one of the motivations behind the wave of renovations now sweeping the world's museums. The other, of course, is financial. Museums must attract paying customers in order to remain viable: these days, they are in the leisure business, competing with theme parks, cinemas and other manifestations of popular culture.

The best displays now combine superb specimens with attractive design, interactive technology, graphics and text, leading the visitor on a preplanned course through the subject. The new Earth Galleries at the Natural History Museum in London provide one instance: where once there were racks of rocks, each labelled with incomprehensible terminology in very small print, there is now a spectacle that entices one to explore the Earth's surface and interior. The revamped fossil halls at the American Museum of Natural History in New York provide another example. There, fossils are arranged along a phylogenetic theme, and the visitor is drawn along the family tree of the vertebrates. Specimens are splendidly displayed — and, as it happens, in a way that reflects the research strengths of the museum.

High on the agenda of the projected restoration in Paris should be the reunion (after a long and unhappy separation) of evolution and anthropology. In the 1960s, many anthropologists shied away

from evolution for fear of being branded racist. Their discipline was thus denied access to a vital perspective of its main substrate, human diversity.

That tide is turning, and the Musée de l'Homme would be well placed to benefit. It would be the ideal forum in which to present recent results from the Human Genome Project, the initiatives to appreciate human genetic diversity, the connections between ethnic, cultural and linguistic diversity, the pressures of population on the environment, the contributions of ethnobotany and folk medicine to modern pharmacology, the latest fossil finds, and much more besides. Art objects too have their place within such a scheme, but as specimens to punctuate a connected narrative display, not just as ends in themselves. □

Moving targets

Migrant researchers with a shadowed past are a growing problem for research institutions.

SCIENTIFIC fraud and malpractice can be a traumatic experience for all involved. This is most obvious for those whose scientific and personal judgement has been exposed to harsh scrutiny, whether in public or in private. But it also applies to others who may have been involved in establishing whether or not a deliberate act of falsification or deception has taken place — a procedure with which those trained in legal rather than academic disciplines are inevitably more comfortable. Understandably, the individuals and institutions alike may seek to minimize both discomfort and risks by keeping quiet.

What happens when a researcher over whom some professional shadow has fallen wishes to move elsewhere? The employers-to-be have their own need to know the background, yet it is of primary importance that the interests of those accused of scientific misbehaviour receive adequate protection. Reckless charges of misconduct can wreck a promising career, and may be prompted by professional jealousies.

Nevertheless, well-founded doubts about individuals should be passed in confidence from one institution to another. Yet obstacles can arise, particularly at the international level, where information flow is often reduced to formal communication, geographical distance may reduce a sense of mutual responsibility, informal contact is at a minimum — if it exists at all — and fears of expensive and time-consuming litigation can dominate all other considerations.

Unfortunately, such cases may now be affecting an increasing number of the world's major research institutions (see pages 107–108). In response, it is tempting to dream up some institutionalized means of addressing the problem. Learned societies and academies might wish to collaborate in examining the issue, but the practical and legal difficulties confronting any formal procedures would be severe, to put it mildly. Those seeking to employ researchers from overseas have no alternative but to bear this phenomenon in mind when, ever more carefully, they seek information regarding a candidate's suitability. □



International recruitment highlights need to track scientific behaviour

San Diego. The increasingly tough stand being taken by US institutions against those committing misconduct in research is creating a new group of scientific fugitives — researchers who seek posts in laboratories abroad after investigations by American institutions cast a cloud over their careers.

Thus a number of scientists alleged to have fabricated research data, to have falsified their academic credentials, or suspected of sabotaging experiments of their colleagues, have moved to research posts in Europe, after inquiries into their activities in the United States.

But this phenomenon has raised questions about how research institutions can obtain clear information about the history of researchers they intend to appoint, particularly if such researchers move to countries that lack a well-defined system for dealing with allegations of misconduct.

"This is a serious problem around the world," says Martin Raff, professor of biology at University College London, adding that there tends to be a lack of communication between institutions on fraud problems. "These cases keep occurring, but there seems to be no good way of dealing with them. The problem is a fear of litigation. Something needs to be done."

Stephen P. Lock, a former editor of the *British Medical Journal* and now a research fellow at the Wellcome Institute for the History of Medicine in London, describes the situation as a "a storm signal".

In the era of faxes and electronic mail, Lock points out, there is no excuse for not checking a scientist's credentials. But, in the face of a growing number of civil lawsuits in individual countries, international monitoring faces considerable challenges.

A typical case involves Kimon J. Angelides, a 44-year-old neurobiologist, who moved to the University of Durham in Britain last year, after losing his post at Baylor College of Medicine in Houston, Texas, for allegedly fabricating research data. Last August, Angelides filed lawsuits in state and federal courts in Houston against Baylor and the key officials in the university's investigation. Angelides is suing for defamation, blacklisting, wrongful termination and breach of contract, claiming that he had brought more than \$7 million in research grants to Baylor.

In his lawsuit, Angelides states that Baylor terminated his contract in March 1995 after the university's investigating committee had found that he had falsified or fabricated data in five scientific articles

and in grant applications to the US National Institutes of Health (NIH). After the verdict, Angelides, who had been at Baylor since 1986 and was a tenured professor in the departments of cell biology and biochemistry, was escorted from his laboratory by university security guards.

Baylor subsequently notified the NIH of the findings of its investigation, prompting the agency's Office of Research Integrity (ORI) to open an inquiry. ORI officials decline to comment, as the agency has not reached a conclusion on the case. Sam Crocker, Baylor's counsel, also declines to comment, as does Angelides, who is now a professor of cell biology in the department of biological sciences at Durham.

James V. Pianelli, Angelides' attorney, describes Baylor's procedures as flawed. "There was denial of due process," he says. "Allegations were made by folks with axes to grind." A trial date of February 1997 has been set for Angelides' civil lawsuit against Baylor. ORI has put its investigation in abeyance until after the trial, says Pianelli.

Asked what Durham knew about the allegations against Angelides, Evelyn A.V. Ebsworth, the university's vice-chancellor, said: "I am not aware of the details of the problems. I have discussed some difficulties at Baylor."

Earlier in his career, Angelides fell foul of

officials at McGill University in Montreal, Canada. In 1980, officials say, he was found to have listed false academic credentials in a *curriculum vitae*, which also included misleading citations for research articles that were not published.

Rose M. Johnstone, then chairwoman of McGill's biochemistry department, says she discovered by "a fluke" that Angelides had padded his CV with four or five unpublished articles and falsely listed an undergraduate degree from Harvard University. Angelides received his undergraduate degree from Lawrence University in Wisconsin. His doctorate in chemistry came from the University of California at Santa Cruz.

Johnstone says that she requested Angelides to resign, adding that she never wrote him a letter of recommendation. "He said I was making a terrible mistake because he would have brought notoriety to the department," Johnstone says. "I said that was one thing we didn't need."

Pianelli denies that Angelides misrepresented his credentials or research at McGill. Angelides left Canada for the University of Florida's College of Medicine in Gainesville, where officials say he was never investigated or found to have engaged in impropriety. But there were questions about his research.

"He was someone who could dazzle you with an enormous amount of research," ►

Rotation experiment returns to Cologne

Munich. **A famous experiment to prove that the Earth spins on its axis, originally conducted in 1852, was recently repeated in the southern transept of Cologne Cathedral, Germany's largest cathedral (right), to celebrate the 200th anniversary of the birth of the physicist Caspar Garthe.**

Garthe suspended a lead-filled brass bowl on a steel wire from the cathedral's 45-metre-high chancel, and used two wedge-shaped scales placed just above the floor to measure its deviation from its plane of swing, a proof of the Earth's rotation.

The admission fee of 20 silver groschen, paid by 1,000 visitors during one month of daily performances, was used to help construct the Gothic cathedral. In 1851, a year before Garthe, the French physicist Léon Foucault more famously repeated the experiment in the Panthéon in Paris. □

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Boecker

► says a prominent Florida scientist. "But senior faculty raised eyebrows that his work was not well replicated," he says.

The separate case of Timothy J. Baldwin shows the difficulty of uncovering problems that may follow researchers moving internationally between scientific institutions. After receiving his doctorate in biology in 1985 from University College London, Baldwin went to the Massachusetts Institute of Technology (MIT) as a postdoctoral fellow in the NIH-funded laboratory of Steven J. Burden, a neuroscientist.

Whilst at MIT, those working in Burden's laboratory, including Baldwin, were investigated following sabotage to experiments, but no-one was found responsible. Baldwin left in August 1989 for a postdoctoral position at the University of California in San Francisco (UCSF). About two months after he left Massachusetts, MIT records say, Baldwin was asked to return because some molecular biology experiments he had conducted for an article published with Burden (see *Nature* **341**, 716; 1989) could not be reproduced.

Baldwin went back to MIT, attempted unsuccessfully to repeat the experiments one day and then admitted himself to a mental hospital, MIT records say. After leaving hospital, Baldwin returned to research at UCSF, where officials were unaware of what had unfolded in Boston.

These events prompted Burden to retract a major part of the published study (see *Nature* **345**, 364; 1990) and caused MIT to initiate an investigation into possible misconduct. This probe became a contentious affair between Baldwin and Burden. MIT decided in 1992 that it could not conclusively establish whether Baldwin had falsified the research results, largely as his laboratory notebooks had disappeared.

NIH officials reached the same conclusion. But they added in a separate report that Burden did not engage in scientific misconduct and acted properly in exposing

the claims by Baldwin that could not be supported. The NIH report also expressed reservations about the fairness and objectivity of MIT's handling of Burden's complaint.

Kenneth D. Campbell, a spokesman for MIT, insists that Burden was treated fairly. "There was no bias against Burden at MIT, as asserted by the OSI [the predecessor of ORI]," he says. Burden subsequently left MIT for New York University.

Problems with the sabotage of experiments also occurred at UCSF in the laboratory of Lily J. Jan, where Baldwin worked for four years after leaving MIT. Baldwin was investigated with others by UCSF at the time, as he was one of those working in the laboratory, but the investigation did not find who was responsible.

After he left UCSF early in 1994, Baldwin returned to the United Kingdom, taking up a position in the experimental pathology division at Guy's Hospital in London. But problems arose there. Baldwin acknowledges resigning his research position at Guy's in September 1995 while he was under investigation for harassing a colleague by sending him offensive sexual material.

Guy's Hospital spokeswoman Carol Meads declines to discuss Baldwin's case, except to note that the hospital has policies for checking the references of all applicants. Baldwin says he resigned from Guy's during the investigation because "it was obvious I was not getting a fair shake". He acknowledges he sent the offensive material, but declines further comment.

At Glaxo Wellcome's Geneva Biomedical Research Institute, the director, Jonathan K. C. Knowles, says he knew that he was hiring a scientist with a troubled past when he employed Annemarie Surprenant, a neurophysiologist, in late 1993.

Officials at Oregon Health Sciences University in Portland had found in 1992 that Surprenant falsely claimed in three grant applications to the NIH that she earned a medical degree in 1976 from the University



of Illinois at Chicago. There was no record of her having attended that university. She received her doctorate in physiology in 1979 from Monash University in Australia.

It was learned during the Oregon probe that Surprenant had falsely stated that she had a medical degree when she was hired for a job at MIT in 1983. The Oregon report says Surprenant was recruited to MIT by R. Alan North, a British-born neuroscientist who is now her husband.

In May 1994, Surprenant acknowledged scientific misconduct in an ORI settlement agreement, which barred her from receiving US government grants for three years. In an interview, she at first said the false listing of a medical degree on her CV was "a clerical error". Then she acknowledged that she "had made a mistake".

"I thought I had a medical degree," said Surprenant. But, she added, the difficulties in the US are irrelevant. "I signed the ORI settlement to put this away, so my research could continue. I stand behind every piece of research I have ever done."

Knowles says Surprenant and North moved as a research team to Geneva, and he has been extremely happy with their work. "My job is not to supervise the scientific establishment of the United States," Knowles says. "My job is to maintain a high standard of science at this institute." He argues the NIH acted properly and well in Surprenant's case. "Obviously, mistakes were made. I was interested in having (Surprenant) do another job. She does it well."

Whatever the circumstances surrounding individual cases, an increasing number of scientists feel that more effective procedures are needed by which scientific institutions can communicate freely — and in confidence — about the past behaviour of researchers without the risk of litigation.

"Sooner or later most large institutions have to deal with this kind of problem," says Raff. "There are powerful forces on individuals to keep this type of thing quiet," he says, arguing that it is not in the interests of research group leaders, for example, to publicize misconduct that has occurred in their own laboratory. "But it cannot go on like this. We need ways to protect institutions and scientists."

Rex Dalton

Private investigators track down fraudsters

London. A former British police detective and a medical doctor have turned 'fraud-busters' and set up a private agency to investigate fraud and misconduct in clinical research, particularly that related to the testing of drugs.

Medico-Legal Investigations is offering to conduct investigations into possible research and other fraud in health service trusts, health authorities and the pharmaceutical industry. It is run by Frank Wells, former director of medical affairs for the Association of the British Pharmaceutical Industry, and Peter Jay, a former police detective.

Jay has since been investigating fraud and other malpractice for the solicitors to the General Medical and Dental Councils, while Wells is the author of a text-

book on medical research fraud. The pair claim in their publicity brochure to be able to achieve "faster results" than more conventional procedures, to defuse "delicate situations" and to turn around "adverse press criticism".

But the prospect of latter-day Sherlock Holmeses using magnifying glasses to examine laboratory notebooks has been met with some nervousness in the medical community. David London, for example, registrar of the Royal College of Physicians of London, says that he is cautious about the idea, which has not yet been discussed at the college. He says that he would prefer to see the agency brought under "the umbrella of an official body, rather than [being] freelance". □

Anthropology museum fights art gallery plan

Paris. The fate of the world's largest anthropology museum, le Musée de l'Homme (Museum of Mankind) in Paris, is hanging on the outcome of a bitter dispute between museum scientists, who want to renovate the existing museum, and art specialists who want to turn it into a museum of primitive, or 'early', art.

A facelift for the museum has been on the cards since 1989, when François Mitterrand, the Socialist president at the time, launched a FF1-billion (\$200-million) programme to renovate the four museums which belong to the education ministry.

Staff at the museum want to repeat the success of the FF500-million renovation of the Natural History Museum's Grande Galerie d'Evolution, which reopened in 1993 after being closed to the public for almost thirty years. Its exhibitions, complete with light and sound effects, now rank among Paris's top attractions.

But plans to renovate the anthropology museum have been complicated by the government's decision to add collections from the Museum of African and Oceanic Arts. Both collections contain valuable items of primitive art, and in January, Jacques Chirac, the French president, set up a commission to review how the joint collections could be best exhibited. (Chirac, a connoisseur of primitive art, is said to have been influenced in his decision by his friend Jacques Kerchache, an art collector and former dealer in African art.)

The commission is split between two competing visions of the museum. Henry de Lumley — who is director of the Museum of Mankind's parent body, the Natural History Museum, and a member of the commission — has proposed a FF400-million renovation that would expand the museum's research capacity and provide a spectacular

public showcase for some of its 735,000 objects, with exhibitions showing the sweep of mankind's social, cultural and biological evolution. The remainder of the collections would be housed in a purpose-built 6,000-square-metre facility beneath the museum. (Many objects are at present stacked haphazardly in rooms and corridors at the museum.)

But several members of the commission are said to favour turning the museum into a

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Science or art? De Lumley (right) wants to expand the research capacity of the Paris Musée de l'Homme (above).

centre of primitive art; research would be abandoned, and large parts of its collections might be sold. Members of the commission said to support the creation of a museum of primitive art include Kerchache, Françoise Cachin, the head of the French museums authority, and Jean-Hubert Martin, the head of the Museum of African and Oceanic Arts.

De Lumley describes the choice between his proposal and the latter as "between renovating or dismantling" the museum. To defend his proposal, De Lumley has broken the requirement that members of the commission refrain from commenting publicly on its deliberations, and has mounted an international campaign in support of his

proposal to create a "Grand Musée d'Histoire Naturelle de l'Homme".

Those supporting De Lumley's proposal include two Nobel prizewinners, François Jacob and Jean Dausset, as well as Ellen Futter, president of the American Museum of Natural History in New York, and Theya Molleson, a member of the department of palaeontology at the Natural History Museum in London.

Jacob says the proposal to turn the museum into a centre of primitive art would be a "regrettable error". Torben Lundbaeck, director of the Danish National Museum of Ethnography, says that the museum's distinctive characteristic has been to nurture links between biological anthropology, archaeology and ethnology, as

well as between research and teaching. "A change in these relationships would seriously damage the identity of the museum," he says.

The commission recently submitted its report to Chirac, who later this year will inform the government of his preferred option. According to Christine Albanet, his adviser for education and culture, the probable outcome will be a compromise, resulting in the creation of a new museum of "civilization" that would cater both for science, in particular ethnology, and for art. "I think the president's desire is that everybody will find an advantage," she says.

But the new museum would be separated from the Natural History Museum, and belong not only to the ministry of education but also to the ministry of culture. The Natural History Museum describes this arrangement as "unacceptable", and claims that the proposal itself is nothing more than a fudge aimed at disguising the "destruction" of the Museum of Mankind.

Declan Butler

Labor resists reductions in Australian research incentives

Sydney. Australia's opposition Labor Party is to challenge several of the recent budget proposals for reducing expenditure on research and higher education by the new Coalition government in a move that, unusually, has placed science and university policies centre-stage in Australian politics.

In particular, Kim Beazley, the leader of the opposition, is to oppose reductions in tax incentives for industrial research and development (R&D) worth A\$1.9 billion (US\$1.5 billion) over four years, and a proposed doubling of deferred course fees, known as higher education credits (HECs), for students choosing science-based courses, worth A\$1.13 billion (see *Nature* 382, 659 & 383, 8; 1996).

Both proposals are vulnerable, as they require amendments to tax laws which must pass the upper house, the Senate, where the Liberal and National parties are a minority. But the opposition party does not intend to challenge the proposed pruning of operating grants for universities and research agencies. These are integral parts of the budget bill, defeat of which would lead to a new election, which the weakened opposition is not pursuing.

In a television interview, Peter McGauran, the Science and Technology Minister, described the opposition's move as "grossly irresponsible". He claimed that the reduced concession of 125 per cent is "a very generous tax incentive" and "the 150 per

cent was very inefficient". But Simon Crean, the opposition spokesman on industry, predicted that R&D "will go overseas as a result of the cuts".

The exchanges on the budget have overshadowed an announcement by McGauran of a year-long inquiry to develop a strategy for marine research to support Australia's recently increased 200-km economic zone, among the largest in the world. Meanwhile, the vice-chancellors of the Universities of Sydney and Melbourne have called on the Senate to approve the HECs bill to protect higher education against further financial pressures.

Peter Pockley

• See the Commentary, 'New era of fiscal restraint in Australia', page 117

Singapore plans to double research cash

Singapore. The government of Singapore last week announced ambitious plans to double the spending of its National Science and Technology Board (NSTB) to S\$4 billion (US\$2.9 billion) over the next five years.

The move is expected to sustain the rapid growth in both industrial and government research and development (R&D) in the island state. But, unlike the preceding five-year plan, which focused on industrial research and technology, 30 per cent of the total has been earmarked for long-term strategic research, a move from which Singapore's universities stand to benefit.

NSTB was set up in 1991 under the Ministry of Trade and Industry to promote 'industry-driven' R&D under a five-year plan aimed at catching up with advanced nations in fields such as electronics, biotechnology, information, materials and manufacturing technology. A total of S\$2 billion was set aside for the five years, of which S\$1.1

billion has already been spent and the rest is committed, for example, in new buildings and salaries for researchers at new institutes that are in the process of opening.

The money spent on research institutes and centres at the National University of Singapore (NUS) and Nanyang Technological University, has already been substantial. Among them are the Institute of Molecular and Cell Biology (IMCB), the Institute of Microelectronics, the Institute of Molecular Agrobiotechnology and the soon-to-be-established Institute of Materials Research and Engineering, all at NUS.

Investment in the university sector is expected to grow significantly in what is now called the National Science and Technology Plan, rather than just the technology plan. Bernard Tan, dean of science at NUS, says he is "delighted" with the plan, and expects more funds for 'strategic upstream R&D' in which the universities can participate.

The results of the first five years have been spectacular. Gross expenditure on R&D by government and industry has leapt from S\$572 million in 1990 to S\$1.33 billion in 1995, almost two-thirds of the latter being contributed by industry. Over the same period, the number of researchers and engineers in Singapore has almost doubled, from 28 to 45 per 10,000 of population, bringing Singapore in line with some advanced European countries.

The government has nevertheless fallen far short of a 1991 target of raising national spending on R&D to 2 per cent of gross domestic product by 1995. In fact, this figure has levelled out at about 1.1 per cent since 1992. According to Teo Ming Kian, the

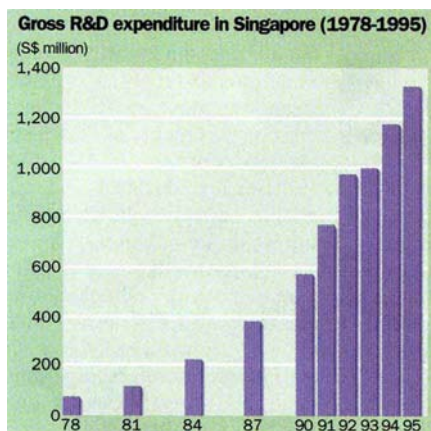
chairman of NSTB, announcing the new plan last week, this is partly due to faster than expected growth of the Singapore economy. He pointed out, however, that the government's grants of S\$735 million have brought S\$2.7 billion of R&D investments by industry, a multiplier of 3.6.

"We are no longer a low-cost production base," said Yeo Cheow Tong, the minister for trade and industry, at the announcement. "We will increasingly face a 'competition squeeze' from developing countries on the one hand and the developed nations on the other," he said, adding that Singapore had to become an increasingly technology-intensive economy.

One of the biggest problems facing Singapore is a lack of manpower. Local graduates in science and engineering still tend to choose financially rewarding careers in local companies rather than pursue postgraduate research. Most of the rapidly growing population of science and engineering graduate students in Singapore is being recruited overseas, particularly from India and the Chinese mainland (see *Nature* 383, 13; 1996).

The plan calls for the number of research scientists and engineers to be increased from the present level of nearly 8,000 by 5,000 over the next five years, and 7,000 more in the following five years. To help in this process, 20 "world-class R&D leaders" will be recruited to act as 'magnets' to attract good local and foreign talent, an approach that has been used with success at several of NUS's new institutes. By 2000, between 12 and 15 new 'top-class' university laboratories will be established at Singapore's two universities.

David Swinbanks



Strategy seeks to make the most of Scottish science

Edinburgh. The Royal Society of Edinburgh, Scotland's leading scientific organization, has joined forces with Scottish Enterprise, which is backed by the UK government, to launch a long-term strategy designed to encourage the commercial development in Scotland of scientific discoveries made in its universities.

According to officials responsible for the initiative, it is based on an awareness that, despite Scotland's success in attracting foreign-owned manufacturing plants in fields such as microelectronics, its own industry has been slow to capitalize on the scientific output of its 13 universities.

Measured in terms of scientific papers per head of population, Scotland ranks third in the world (after Israel and Sweden), and although it has only 9 per cent of the UK population, it boasts 15 per cent of the UK academic science base. Yet relatively few of its scientific achievements are exploited by Scottish industry.

Drawing on the experience of 'networking' initiatives, both privately and publicly funded, in other countries — particularly the United States — the goal of the initiative is to build a coalition between the academic world, business and finance groups, and the public sector.

"The short-term job impact may not be large," says Crawford Beveridge, the chief executive of Scottish Enterprise. "However, success will help determine the future competitiveness and well-being of the Scottish economy."

In launching the new initiative at Heriot-Watt University outside Edinburgh — which claims to be the site of Europe's first science park — George Kynoch, the minister for industry at the Scottish Office, said that part of its aim is to complement the Technology Foresight programme introduced by the government in its white paper of May 1993. "Technology Foresight has been about identifying opportunities,"

said Kynoch. "The strategy that we are launching today is about making sure that the commercial potential of that research is realized."

The initiative, known as Technology Ventures, will be promoted by a high-level 'leadership group' being set up by Scottish Enterprise, and including both industrial and academic members. The Royal Society of Edinburgh has been the joint sponsor of a two-year programme leading to the initiative, and is keen to see it succeed.

"With over 1,000 fellows from all academic disciplines, we can bring to bear a range of scientific expertise," says Tom Johnston, the president of the society and a former vice-chancellor of Heriot-Watt. "This is the first occasion on which Scotland's principal economic development agency and her leading learned society have come together to work on a theme in which both have a major interest and the expertise to pursue it."

David Dickson

Labour party pledges to end privatization of UK laboratories

London. Britain's opposition Labour party promised this week that if — as widely tipped — it wins next year's general election, it will halt the programme launched by the Conservative government under which all government laboratories are being systematically evaluated as potential candidates for privatization.

The programme, known as the Prior Options Review, has already led to responsibility for a number of government laboratories being passed to private-sector organizations under a variety of arrangements, ranging from direct sell-offs to the contracting-out of managerial responsibility. Included in the current review are many laboratories belonging to the six research councils.

But such moves have been highly controversial among those — including the labour unions representing the scientists and technicians working in such institutions — who believe the government should maintain central responsibility for them.

In a policy statement on industrial strategy issued this week, the Labour party describes the laboratories as a "vital national resource and a source of crucial research expertise". It adds that they also have "a vital role to play in offering independent, impartial advice to government".

The document, which was launched by Adam Ingram, the party's shadow minister for science and technology, during the annual meeting of the British Association for the Advancement of Science, promises that a future Labour government will take steps to increase the involvement of women in science. It will also try to "strengthen and rationalize the current UK intellectual property rights system" — although no details are given of how this might be done.

But there is no indication that the Labour party is prepared to commit itself to providing extra funds to support research, either in universities or industry. And the party says that, even though it sympathizes with scientists who were "dismayed" at last year's decision to move the Office of Science and Technology into the Department of Trade and Industry, with no prior consultation, it has no plans to reverse the move.

"Another process of rapid change would not necessarily be in the best interests of the scientific community," says the document, which sets out the party's general approach to the area of innovation, design, science and technology. It merely promises that a Labour government would "undertake a thorough review of the effectiveness of the new structure, in consultation with industry and the research and academic sectors".

David Dickson

ESA members reluctant to replace Cluster mission

Munich. Prospects are looking bleak for rebuilding and relaunching the Cluster mission of the European Space Agency (ESA), whose four satellites were lost when Ariane-5 exploded during its first test flight earlier this summer. A search for financing from European governments has so far drawn a blank.

Nevertheless, scientists involved in the mission reaffirmed their determination last week, at a meeting of ESA's Solar System working group in Paris, to continue the fight for a full recovery plan. "We're not going to give up the battle before it has begun," says André Balogh of Imperial College, London.

Cluster was intended to study the electrical and magnetic field environment of the Earth. It was scientifically unique, as its four-satellite configuration would have allowed the first three-dimensional analysis of this environment.

After the accident, ESA drew up a proposed rescue plan that suggested launching Cluster's single spare satellite, called Phoenix, on the next Ariane-5 test launch early next year. ESA suggested that this could be joined three years later either by three further identical satellites, or by three minisatellites, to allow multipoint measurements (see *Nature* 382, 102; 1996).

Like the first test launch, this would have been free. But this idea has now been abandoned by ESA's science directorate, headed by Roger Bonnet, because of fears that the problems that led to the first failure may not be rectified in time.

Temper in ESA have become severely frayed over the issue. Jean-Marie Luton, ESA's director general, is reported to be unhappy that such criticisms have reached the public arena, and is making no public comment on the Cluster affair.

The science directorate, in turn, is

unhappy that it will not be offered a discount to fly Phoenix on the first commercial launch of Ariane-5, due at the end of next year. The full price is around ECU130 million (US\$165 million). Bonnet, with the support of the scientific community, is trying to achieve a compromise from Luton and ESA's launcher division.

The two teams set up to cost the additional satellite options learned last week that prospects for financing are not good. So far no money has been found within ESA to finance the full-sized units, and member states — which, it was hoped, would pay for the minisatellites — have not indicated a willingness to foot the bill.

But discussions are continuing, and these options will remain open until the Science Policy Committee, ESA's decision-making body comprising representatives of member states, meets again at the end of November. Meanwhile ESA is negotiating with industry to try to reduce costings for either option.

It has been an unhappy summer for solar scientists. A Russian-led mission, Interbal, which shares some of Cluster's scientific aims, suffered a major setback last week when the power system of the smaller of its two satellites failed after launch. As a result, scientists will lose the opportunity offered by the mission of measuring the magnetosphere contemporaneously at two points.

Meanwhile, the scientists involved in Cluster are demanding some financial compensation from ESA for the Ariane accident, perhaps in the form of a free or discounted launch, along the lines being discussed. They are angered by the conclusion of the independent board of inquiry that management negligence contributed to the software failure which sent the rocket off course and required it to be destroyed (see *Nature* 382, 386; 1996).

Alison Abbott

Nuclear treaty text set for UN approval

Paris. Australia led an attempt this week to have the current text of the Comprehensive Test Ban Treaty (CTBT) sponsored as a resolution to the UN General Assembly. As 121 of the 185 UN countries have agreed to co-sponsor the resolution, it is virtually assured the two-thirds majority it needs to pass.

The co-sponsors include the five nuclear weapons states: the United States, Russia, China, France and the United Kingdom. Of the three threshold weapons states, Israel is sponsoring the Australian resolution, while Pakistan is said to be likely to do so. India, which last month vetoed the adoption of the

text by the usual route — consensus at the Conference on Disarmament in Geneva — still refuses to sign, complaining that the text provides insufficient commitment to nuclear disarmament.

Countries that co-sponsor the text agree not only to support the treaty but also to oppose any amendments and to seek further co-sponsors. This means that the 121 co-sponsors are already more than enough to block amendments, given that these require a two-thirds majority to pass. Supporters of the text fear that any amendments could upset hard-won consensus on other contentious issues. □

'Petrol from plants' claim baffles Indian scientists

New Delhi. Indian scientists have been baffled by the claims of a 30-year-old self-taught chemist from the southern state of Tamilnadu that he is able to power a scooter using 'petrol' produced by adding leaves and bark extracts from a native herb to tap water.

Ramar Pillai demonstrated his backyard experiments to a scientific audience last week in the chemistry laboratory of the Indian Institute of Technology (IIT) in New Delhi. Scientists who watched the water turn into an yellowish, inflammable liquid admit that they have no explanation at present.

But they remain convinced by the demonstration. "It is just unbelievable but true," says N. K. Jha, head of the chemistry department. Valangiman Ramamurti, the secretary of the Department of Science and Technology (DST), who himself performed the experiment said: "Like everyone else I was sceptical, but we now know it is not an Indian rope trick."

Pillai's curiosity had initially been stimulated as a schoolboy when a spark from a camp-stove ignited the leaves of a distant plant during a school picnic. Ten years later, he tracked down the 'burning plant' and started experimenting with it, using knowledge picked up from chemistry textbooks.

The process he developed to make the fuel involves cooking the leaves of this plant — although other parts will do — for about ten minutes in hot water to which a little salt and a few drops of lemon juice or citric acid have been added. As the liquid cools, certain chemicals — believed to be catalysts — are added, and the contents allowed to settle.

Because it is lighter than water, the herbal fuel rises to the top, and is filtered off. Pillai says his fuel would cost one rupee (3 US cents) a litre. Commercial production should not be a problem, because the plants involved grow like tea bushes, with leaves continuing to grow to replace those that have been plucked.

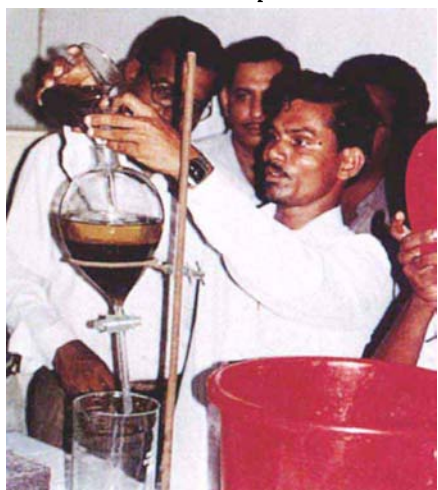
Last week's experiment started with 55 grams of the herbal leaves and bark, and 1,000 millilitres of tap water. Just over half an hour later, this had produced 460 millilitres of a liquid that burned and smelled like kerosene, and on distillation yielded a fraction boiling at 170°C).

According to the National Chemical Laboratory in Pune, analysis of the fuel has shown it to be a pure hydrocarbon, although no details have been published for fear of pre-mechanical patent applications. Tests at the mechanical engineering department of IIT Madras found the fuel to be "better than [real] petrol" in terms of economy. Furthermore, being of plant origin, it contains no

sulphur and the resulting exhaust is therefore cleaner.

"We have no doubt that we are sitting on something very big, but we want to proceed cautiously and systematically," says Ramamurti. Officials at the DST strongly believe that the herbal process, if commercialized, could have a major impact on the Indian economy.

The DST has already agreed to provide funds for a pilot plant at Rajapalayam, capable of producing 300 litres a day. At the same time, India's leading chemists will try to understand how the process works, as



Liquid gold? Pillai produces his 'fuel' using leaves of a native herb and water.

plans are drawn up for a production capacity of one million litres a day. According to Ramamurti, all of India's main scientific agencies will be provided with the necessary funding to participate.

Scientists who have watched the experiment accept that the process generates a hydrocarbon fuel. But they describe this as a 'miracle' in the absence of accepted experimental protocols, and lack of details of the herb and catalysts used.

Two factors in particular remain baffling. "The process is too fast," says Ratna Choudhury, a chemist at IIT who has been studying methods of extracting hydrocarbons from the so-called 'petro-crops' such as *jojoba*. Second, although the hydrogen could, in principle, be provided by the water, they have no explanation for the presence of large quantities of carbon; a suggestion that it might be 'sucked in' from the air is widely considered to be implausible.

Pillai says that he will disclose the secrets only after his invention is patented. DST hopes that the patenting process will be completed within two months, at which point controlled experiments will be able to start.

K. S. Jayaraman

Foundations funding biomedical bodies 'should shift focus'

San Francisco. Private foundations should play a greater role in helping academic biomedical institutions adjust to the tightening of federal funds, increasing financial pressures on medical centres, and the growing role of industry in research, according to the University of California's San Francisco Center for the Health Professions. A report prepared jointly with the Pew Biomedical Scholars Program warns that foundations are exacerbating a mismatch between training and job opportunities by continuing to emphasize doctoral training and fellowships.

Carolyn Asbury, director of the health and human services programme of the Pew Charitable Trusts, says that foundations should instead be trying to meet industry's demand for more interdisciplinary training, to focus on policy issues, and to develop centres where investigators from different backgrounds can share resources.

The report predicts that the 50-per-cent increase in doctoral awards in the biological sciences over the past 10 years will lead to an oversupply of biomedical researchers.

Funding tends to support established laboratories as research becomes more costly and the emphasis on practical applications grows. This further harms the prospects for promising young scientists, says Edward O'Neil, co-director of the Center for the Health Professions and principal author of the report.

Interviews with 200 Pew scholars indicate that junior-level faculty members are growing disenchanted, with good scientists leaving for business, law and other fields. The scientists interviewed by O'Neil were worried they might not be able to continue in science, even though they had received one of the most sought-after young faculty awards in the United States.

Federal spending on biomedical research has not kept pace with increased costs, the report points out, in spite of a 6.5-per-cent increase in federal funding for the National Institutes of Health (NIH) during the 1996-97 budget cycle. In 1992, industry started to become the most important source of money for biomedical research. In 1994, industry spent \$17.1 billion on such research, compared to \$14.5 billion from the federal government.

O'Neil says NIH grants and the reward systems at universities must change to encourage smaller laboratories to flourish. Large laboratories should develop a new class of mid-level senior scientists, instead of just a principal investigator plus fellows and graduate students. Universities should seek ways to fund a scientific infrastructure independent of grants. Sally Lehrman

Breast cancer grant policy comes under fire

Washington. Two scientists from Yale University in New Haven, Connecticut, are complaining that they have been denied funding by a huge federally funded breast cancer research programme, even though they received impressively high scores when ranked by scientific peer-review panels.

Their complaints have rekindled a fierce dispute that took place when the Breast Cancer Research Program (BCRP) was set up by Congress in 1993, after intense lobbying by advocates of breast cancer research; the dispute centres on the extent to which peer-review judgements might become subservient to other factors in allocating research grants (see *Nature* 363, 195; 1993).

Proposals from Barry Kacinski, a radiation oncologist and tumour biologist at Yale University School of Medicine, and Trevor Williams, an assistant professor of biology at Yale University, were placed in the 98th and 90th percentiles respectively. But they were rejected in June on the basis of decisions by a second, non-peer-review panel.

Williams' \$860,000 proposal sought to find the relevance of a particular gene to breast cancer and breast development. Kacinski's \$699,000 proposal aimed at developing a non-toxic therapy for advanced breast cancer, based on the role of glucocorticoids in invasion and metastases.

Scientific reviewers described Kacinski's proposal as "extremely well crafted" and "highly relevant to breast cancer research". One called the work "an understudied area that may have a good deal of clinical relevance, especially for patient prognosis".

But although the second 'integration' panel incorporated the scientific review scores in its judgements, it did not rely on them exclusively, adding additional criteria such as the need to avoid redundancy and to support previously underfunded areas.

The manager of the programme, Alan Epstein, wrote last month to Kacinski, whose percentile score was 98.1, that his proposal was "scientifically superb". But he said that this score was only one of many criteria used to assess a proposal.

Indeed, Epstein told Kacinski that 50 per cent of proposals ranked by the scientific peer-review panels as "outstanding" (a score between a perfect 1.0 and 1.5) or "excellent" (a score between 1.5 to 2.0) are not being funded in the current round, for which final decisions will be made before 30 September. Kacinski scored 1.2; Williams scored 1.4. An earlier, unsigned draft of Epstein's letter to Kacinski stated that the average score of a funded grant is 1.6.

Williams wrote in a letter to Representative Rosa DeLauro (Democrat, Connecticut), that the process is "a systematic and wanton abuse of public funding," and that those who petitioned Congress on behalf of women's health issues "are being cheated".

The programme received \$210 million in 1993, and at \$150 million during 1995, the year in dispute. Action by Congress in 1990 forbidding transfers of Defense Department money to domestic budgets prevented the programme being placed in the National Cancer Institute, as its supporters had intended. Instead, it was given to the Army's medical research and material command.

Epstein argues that confidentiality prevents him from commenting on the two proposals under dispute. But he defends the process in general, saying that as more than 2,200 applications were received for about 300 grants, many of the best proposals could not be funded. "Even in that upper tier of outstanding proposals, it was incredibly competitive," he said, calling the scientists' criticisms "unsubstantiated opinion".

Epstein adds that the 24-member integration panel was asked explicitly to shape the direction of funding by considering strategic goals, with authority to reject an application — however strong the science — if it did not meet other goals that included relevance to breast cancer and the potential of resulting in a breakthrough.

Kacinski and Williams complain that the Army is secretive about these other goals. Epstein has told Kacinski that the proceedings of the integration panel in judging his proposal were not recorded, and, in any event, not for public release.

Colonel Irene Rich, the programme's director, said that in future, applicants will receive written records of reviews by the integration panel of their proposals.

But Sara Rockwell, the director of the office of scientific affairs at Yale's School of

Medicine, says that the office will in future advise investigators applying to the BCRP to hedge their bets by applying to other agencies simultaneously. "There appear to be other factors [than scientific merit] of importance here, and we don't know what they are," she says.

Defenders of the programme argue that the integration panel acted precisely as proposed in a report by the Institute of Medicine in 1993. This stated that the fundamental criterion for a successful proposal should be scientific merit, and that "second-rate research should not be supported simply on the grounds of relevance to programmatic goals". But the report also advised that a 'programmatic' review should address goals such as encouraging women and minorities to apply for grants; supporting promising but underfunded research areas; bringing new investigators into the field; and encouraging cross-disciplinary projects.

"We are trying to open up the system and open up the structure," said Fran Visco, a member of the integration panel who is also president of the National Breast Cancer Coalition and a prime mover behind the BCRP. "What this programme is able to do, which many other programmes are not, is a very thorough, programmatic review that makes certain that, regardless of scientific merit, programmatic merit also exists."

Visco said she is "livid" and that it is "appalling" that the scientists are complaining. "They have to understand that just because they're used to getting funded, that doesn't mean they're going to continue to get funded," she says. **Meredith Wadman**

Radar will investigate magnetic storms



London. A radar system designed to investigate fundamental processes in the Earth's ionized upper atmosphere went on-line from the Arctic circle at the beginning of this month. The 32-metre EISCAT Svalbard Radar (ESR, left) is sited on Svalbard, an archipelago of islands halfway between the north coast of Norway and the geographic North Pole. It will investigate the chain of interactions that allow eruptions on the surface of the Sun to disturb the Earth's magnetic field and ionosphere.

These 'magnetic storms' are understood to be capable of disturbing power distribution networks, global positioning systems, radars and satellites. EISCAT is an international research organization involving seven countries — Finland, France, Germany, Norway, Sweden, the United Kingdom and Japan. □

UK revamps genetic engineering panel to focus on policy

London. Growing confidence in genetic engineering has led to the slimming-down of a UK government advisory committee on the safety of the 'contained use' of genetically modified organisms. Derek Burke, former vice-chancellor of the University of East Anglia, and a member of the outgoing committee, says it has proved "a real British success", but changes had become necessary as it had become concerned more with regulatory matters than with science.

For the past 12 years, the Advisory Committee on Genetic Modification (ACGM) had consisted of 19 relevant experts made up of the scientific and public policy communities. The new ACGM has 13 members, who will now focus on "strategic policy issues".

A separate 12-member subcommittee will report to the main committee on scientific and technical matters. Alison Spalding, secretary of the ACGM, says science and policy had been split up to achieve greater efficiency and to focus appropriate expertise on science and regulatory matters.

Kay Davies, professor of genetics at the University of Oxford (pictured left), is the ACGM's new chair. □

James King-Holmes

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Clean-up 'needs more data'

Washington. The US government needs more information on new clean-up technologies, health and environmental effects of nuclear waste, and future land use, before it can decide how best to clean up the Hanford nuclear waste site in Washington state, according to a report from the National Research Council, the operating arm of

the US National Academy of Sciences, which was released this week.

"Instead of choosing one plan now, based on limited information, the federal government should adopt a phased decision strategy," says Thomas Leschine of the University of Washington in Seattle. He suggests delaying some remedial work until there has been more research. According to the report, a recent environmental impact study sponsored jointly by the energy department and the state of Washington considered only the waste inside Hanford's holding tanks, and not the tanks themselves, or waste that has already leaked into the ground. □

DNA test used in murder case

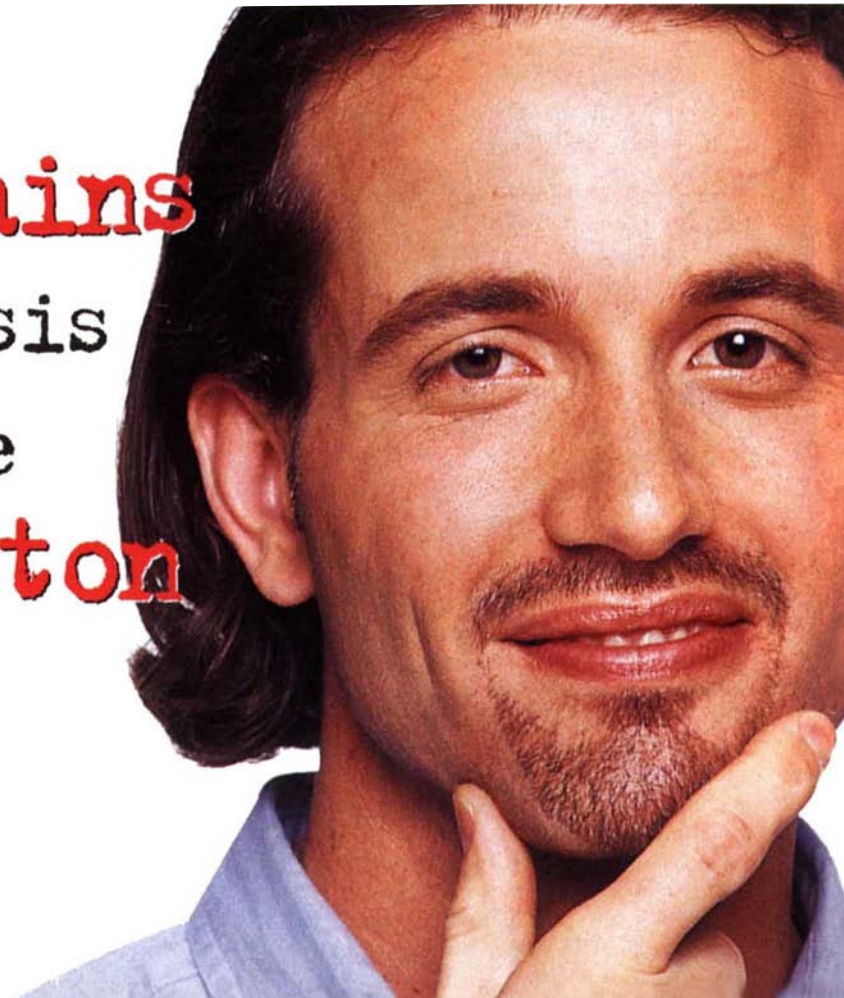
Washington. Evidence from a mitochondrial DNA test has helped to convict a man who raped and then murdered a 4-year-old girl in Chattanooga, Tennessee, in what is reported to be the first use of such a test in legal proceedings. Twenty-seven-year-old Paul Ware was found guilty of the murder of Lindsey Green, killed in 1994. Mitochondrial DNA was used from hair found at the scene of the crime to match that from Ware. Mitochondrial DNA can be detected in hair, bones or teeth; unlike other DNA, it is found in the fluid outside the nucleus of the cell, and carries its own genetic code. □

Scientists back Indian patent bill

New Delhi. India's scientists are lending support to a new bill on patents, which the government proposes to introduce in the parliament session that began earlier this month. Among other things, the legislation will allow patenting of plant varieties, man-made micro-organisms, and gene sequences yielding new products.

A group of leading biologists, meeting recently under the umbrella of the Indian National Science Academy has voiced concern over the delay in amending the 1970 Act, as India has promised to do to qualify for membership of the World Trade Organization.

Doug now stains
electrophoresis
gels with the
push of a button



Lack of product patent is the biggest bottle-neck towards the commercialization of discoveries, says T. C. Anandkumar, chief of a Bangalore-based biotechnology company that has developed four marketable products including a birth-control vaccine for dogs. □

Plan to boost UK mathematics

London. A report co-sponsored by Britain's Engineering Council has proposed ways of reversing the declining mathematical ability of UK pupils, particularly those studying mathematics, science and engineering at universities. One proposal is that potential science and engineering students take school courses identified as essential for a science or engineering university course. *A Mathematical Foundation* notes with concern that some universities have begun to lower entrance requirements for science and engineering courses, while others are altering the course content "so that there is less dependence on mathematics". It calls for more mathematically-qualified entrants to be encouraged into school teaching. □

Award for information on BSE

Washington. The US Federation of American Scientists (FAS) has conferred a unique award on a British scientist for his "outstanding work" in providing information on the outbreak of bovine spongiform encephalopathy (BSE). The award takes the form of a personalized home page on the World-Wide Web (<http://www.fas.org/promed/award.htm>). □

Ralph Blanchard of the Institute of Food Science and Technology in London, a registered charity, has been honoured for up-to-date, balanced and informed views on BSE above the "journalistic frenzy" and "ever more reactive UK governmental actions and ministerial statements". The Excellence in Outbreak Reporting award is conferred by FAS's Pro-Med committee, a project to promote the establishment of a global programme to monitor emerging diseases. □

Winning ideas for space physics

Washington. A low-cost Mercury orbiter, a mission to orbit the Sun using solar sails, and a cosmic-ray experiment on the international space station, are among 19 space physics mission concepts that have been selected by the US National Aeronautics and Space Administration (NASA) for further study over the next two years.

NASA chose the winners from 70 ideas submitted in response to a call for proposals last February. Costs range from \$20 million to \$250 million, and the missions could be ready to start between 1999 and 2005. NASA is making no commitments, however; results of the studies will merely "be considered" in the development of the agency's strategic plan for space science, although they may turn into proposals for flight projects. Short descriptions of the winning concepts can be found on the Internet at http://umbra.nascom.nasa.gov/spd/nmc/nmc_results.html □

Chinese enhance scientists' pay

Hong Kong. A Chinese provincial government is offering the first substantial income supplements to local scientists working for the Chinese Academy of Sciences and the Academy of Engineering, in a bid to make poorly paid jobs in science more attractive. The local government subsidy of 500 yuan a month (about US\$60) will be paid to 43 researchers in the southwestern province of Sichuan, adding significantly to their monthly pay, which is typically as little as between 1,200 and 2,000 yuan a month for scientists in China.

The move follows complaints from scientists that, amid much talk among senior officials about science and technology leading China's growth, poor pay and conditions encourage students to move into more prosperous commercial fields. Other subsidy schemes are already on offer by some individual institutions in China. Among them is extra pay for publication of papers in leading journals, including *Nature* (see *Nature* 378, 541; 1995). □

Are you spending hours at the bench staining gels? You don't have to anymore. Now you can push a few buttons, walk away and return to reproducible staining results.

Hoefer Automated Gel Stainer. the entire staining process is now automated

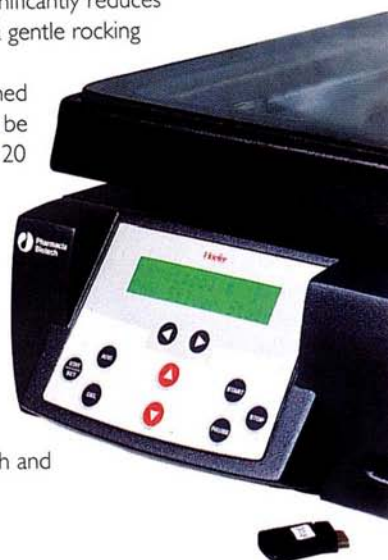
The new Hoefer Automated Gel Stainer makes it happen. It automates reagent delivery and development times to stain your DNA and protein gels reproducibly. It significantly reduces reagent consumption and maximizes mixing while handling your gels with a gentle rocking motion.

The Hoefer Automated Gel Stainer gives you eight pre-programmed protocols for standard silver and Coomassie blue staining—all of which can be modified to your specific needs. What's more, you can design up to 20 protocols of your own and save any of them on a removable "smart key"—keep the key to protect your protocol and simplify your start-up times in the future.

Together with our PlusOne Silver Staining Kits and Coomassie tablets, Pharmacia Biotech can provide you with everything for staining electrophoresis gels automatically. Call us: 1 (800) 526 3593 from the USA; +81 (0)3 3492 6949 from Japan; or +46 (0)18 16 50 11 from Europe and the rest of the world.

Or visit us on the Internet at <http://www.biotech.pharmacia.se>.

Ask for a free brochure; it details how you can save hours at the bench and get reproducible staining results by pushing a few buttons.



"Manual staining is an obsolete concept. It exists no more."

Doug Burtrum,
Immunologist/Research Scientist,
Haiku writer,
New York, NY.

**Pharmacia
Biotech**
Uppsala, Sweden. (And the rest of the world)

READER ENQUIRY NO. 596

The Republican blueprint

SIR — I disagree with your leading article about the treatment of science in the 1996 Republican platform (*Nature* **382**, 655; 1996). Your conclusion that the Republican platform is weak on science is unfair and unfounded. As chairman of the House of Representatives Science Committee and a member of the Platform Committee, I can assure you that the decision of the Republican party was to include in our platform clear language reflecting our commitment to sound science and the role of science in our economy.

Your article leads me to believe that you thought the platform document should be more detailed. Whereas the Democrats pay lip-service to the importance of science with one tiny paragraph in their platform, the Republican party and Bob Dole include science and technology in many parts of our document, demonstrating our realization that science and technology are an integral part of a twenty-first century society, not merely a separate component. The statements made in our platform are reinforced by legislative initiatives in place or under way in Congress. By contrast, the 1996 Democratic platform contains the word 'science' only twice, and does not discuss the importance of basic science at all.

Although the Clinton administration claims to be pro-science, its record shows otherwise. The president does not mention science in his State of the Union addresses. And in his most recent budget, he ignores the importance of science to our economy by opting not to make decisions or prioritize by using plug (notional) numbers for the out-years. The past four years have shown that the Clinton science agenda consists of embracing the status quo and maintaining outdated industrial policies, offering no vision for the future.

The Republican party is different, and the platform document reinforces that difference. Besides devoting an entire section to science, technology and innovation in the twenty-first century, the Republican platform recognizes the benefits of science and technology to improve the standard of living, addresses the need for regulations affecting health, safety and the environment to be based on sound science, and reaffirms our traditional support for medical research through the National Institutes of Health. Our extensive section entitled "A Cleaner, Safer, Healthier America" outlines a positive and proactive agenda based in large part on sound science and technology development. This agenda includes the philosophy that all government environmental decisions are based on "the best peer-reviewed scientific evidence, while encouraging advancements in research...". The Republican commitment to sound science is

reflected in other areas of the document, from defence to energy, with an entire section devoted to space exploration. All of this is lacking in the Democratic platform, despite Clinton's opportunistic statements about increased government support for space in the light of the National Aeronautics and Space Administration discovery about possible life on Mars.

The Republican party platform shows that our party will protect the American technological edge, and encourage and insist on the government's support of the basic scientific research that will result in the new knowledge that is the core of our economic future. Our blueprint for restoring the American dream is bold and visionary — two attributes severely lacking in the Clinton administration's actions or words.

Robert S. Walker

(Chairman)
Committee on Science,
US House of Representatives,
Washington, DC 20515, USA

East and West

SIR — Following a telephone interview, I was quoted (*Nature* **382**, 486; 1996) as having said that "it is hard for a westerner to survive in an east German research institute". I did not say that, or anything else that could have been interpreted in that way. And the observation that follows the quotation, that there remains a clash of cultures and difficulties in establishing competitive research programmes, does not reflect the contents of my statements.

To put things straight, I should like to make the following observations. The situation in east Germany as a whole, and therefore also in east German research institutes, is complicated, regardless of whether an individual comes from the east or the west. My appointment in Berlin was for three years, during which I held positions in Berlin and in Munich. When I had to choose between those two locations, I decided on Munich, where working conditions were better for me. But this does not mean that working conditions in Berlin are unmanageable.

As to the clash of cultures referred to in the last sentence of your article, despite the complicated situation arising from the incorporation of East German research institutes into the research landscape of a unified Germany, there has been a strong spirit of cooperation and a willingness to solve problems.

Peter Russer

Technischen Universität München,
Arcisstrasse 21,
D-80333 München, Germany

Positively discriminating

SIR — Later this month, a conference on "Women in Evolution" will be convened at the University of Arkansas. This "gathering of scientific perspectives" will be sponsored by the Sloan and Rockefeller Foundations, the National Science Foundation, the University of Arkansas and the Arkansas Department of Health and Education.

Only women were invited as guest speakers, as the conference will deal with such feminist issues as nitrogen-fixing bacteria, the evolution of malaria, frog evolution, acoustic behaviour, web-spinning spiders and P-transposable elements in *Drosophila*. The conference will also "chart a course of action that will expand the opportunities for women to participate in science at all educational levels". It is presumably for this purpose that Dr Diana Wheeler from the University of Arizona will deliver a lecture on "Evolution of castes in insect societies".

The logistics of such a meeting are probably mind-boggling. For if the choice of participants is made on the basis of possessing an unpaired X chromosome, then people sharing my single-X disposition, such as individuals with Turner syndrome (XO), will be allowed to participate, while individuals with Klinefelter syndrome (XXY) will be excluded. Alternatively, if the determination is made according to the presence of Y, then the situation for the Turners and Klinefelters will be reversed.

Think what havoc this will cause in the lives of genetic mosaics. The conference will also have an adverse economic effect, as obtaining notarized copies of one's DNA sequence of the testis-determining factor gene may be quite expensive. Interestingly, the sponsoring bodies of the "Women and Evolution" conference refused to consider support for a similarly minded meeting entitled "Bearded Jewish Males with a Slight Limp and Evolution", even though such individuals, who have been "marginalized and excluded" in the past, should also be given the opportunity "to reclaim their power through solidarity and education". The foundations persisted in their refusal even when the title has been changed to something significantly less exclusionary: "Partially Bald, Middle Aged, Pot-Bellied Individuals and Evolution". Given that the tests for baldness, age, and abdominal width are much cheaper than karyotypic tests, I find their response quite inexplicable.

Dan Graur

George S. Wise Faculty
of Life Sciences,
Tel Aviv University,
Ramat Aviv 69978, Israel
e-mail: graur@ccsg.tau.ac.il

New era of fiscal restraint for Australia

Gustav Nossal

The recent first budget of the new Australian Coalition government offers an opportunity to reflect on the way in which Australian science must adapt to an environment of fiscal rectitude.

A FEW weeks ago, Australia's new Liberal-National Coalition government handed down its first budget. The scientific community was facing a situation where the treasurer, Peter Costello, was committed to severe cuts, needing to repair (over two years) a A\$8-billion (US\$6.3-billion) deficit without raising taxes. In this context, Peter McGauran, the minister of science and technology, actually did very well.

The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia's largest research agency, escaped essentially unscathed, although some hoped-for budgetary increases of A\$20 million a year for three years will in fact have to be repaid by asset sales and other savings. Nevertheless, the maintenance of the CSIRO budget was an important victory. The other main government science agencies suffered moderate cuts. Some will compensate by becoming more commercially focused (see *Nature* 383, 8; 1996).

Another triumph was a substantial increase in the funding of the Australian Research Council, the government's main agency for supporting nonmedical research in Australian universities. This increase includes more postgraduate scholarships. A modest increase in funds for the National Health and Medical Research Council for medical research in universities, teaching hospitals and medical research institutes will be matched by new funds from the UK Wellcome Trust. This partnership between public and private sectors is likely to be highly beneficial for medical research.

But these increased funds are dwarfed by government reductions in tax concessions for industrial research and development (R&D). The rather open-ended Research and Development Syndication Scheme has gone, at least partly because it was being used by those more interested in tax minimization than R&D. This is a pity: it would have been fairly straightforward to close the tax loopholes without axing the scheme altogether. Syndication has been replaced by a new "Start" initiative to provide grants, loans and interest-rate subsidies for high-risk research and high R&D spending. Whether this will fill the gap remains to be seen. It is clear, however, that the sums foreseen for Start are much smaller than those saved on syndication. The scheme will also take some time to get going, and in the meantime much momentum will be lost.

The previous, widely praised scheme of

150% tax deduction for industrial R&D has been reduced to 125%. This is unfortunate because largely through this and other government-backed schemes, Australia had begun to redress its traditional imbalance between fundamental and applied research. Industrial R&D has recently been increasing fairly substantially, although from a small base, as have Australian exports based on high technology. To jeopardize this seems imprudent. Some of the more thoughtful business leaders, including John Prescott, managing director of BHP, Australia's largest corporation, have criticized the government for reducing incentive for innovation just as the tide appeared to have turned.

University funding is to be cut by around 5% over the next three years, and student fees awarded through the Higher Education Contribution Scheme are to increase substantially. The implication of this move to 'user pays' is clear, and there is no indication from the government that the extra fees raised should go back into tertiary education. Universities are encouraged to become more independent, with fewer bureaucratic controls. Further, they can retain the fees they charge for students recruited over and above the government-mandated quota in various subject areas. Given that many of these quotas have been difficult to fill, however, it is doubtful that these savings will amount to very much. The offset of a thousand scholarships a year to soften the blow for the best less-well-to-do students is too small even to warrant comment.

Australia's university system seems to be heading in the general direction of that of the United States. There will be more pluralism, strong and weak institutions and extreme competition for student numbers, with recruitment of full fee-paying overseas students an important factor. Invariably, there will be hardship in the academic community, although a freed-up system will certainly reward the most able and sought-after academics and punish the less productive ones. The entry of the marketplace will have to be administered with great care by vice-chancellors, particularly to protect creativity and the right to be different and to dissent.

Fiscal rectitude, not some well-articulated plan for "the clever country", was the key driver of this first budget. The rhetoric in favour of a higher profile for science and engineering is now strong on both sides of the Australian political fence. Even as economically rationalist a body as the Industry

Commission has seen the need for excellence in Australian R&D, praising university research and urging CSIRO to concentrate on the more strategic, pre-competitive aspects of its work. The new growth economics suggests that countries emphasizing an entrepreneurial culture and technological innovation and building a first-class educational system are destined to thrive in the next millennium. So it is sobering to see the proportion of the total pain that will now be borne by the higher-education sector and industry committed to R&D. The 1997-98 budget will probably have a similar flavour.

There are really only two pathways for expansion in fundamental research. The first will be the difficult task of forging a partnership with industry in which the academic sector is valued for what it can contribute, not for short-term tactical problem-solving. The second will be an extension of what has been so well begun through the Cooperative Research Centre Scheme. Initiated by the former Labor prime minister Bob Hawke, the scheme has been strongly supported by the Coalition. It seeks to build collaborative networks (there are already 62) between different fields of science and technology, where academic and government laboratories (particularly CSIRO) and industry work together. A government commitment of (usually) A\$2 million a year must be at least matched (in cash or in kind) by the partners. The result has been a pooling of skills and resources, a new broader commitment to the education of postgraduates, and a better mutual appreciation of problems and challenges. There is a robust tradition of competition in Australian science: it is time to work together more creatively.

Australia's excellence in basic science can continue only if the attitudes of the past can be modified. The traditional handout mentality will no longer work; entrepreneurial self-help will be required. This will include intelligent partnerships with industry and multi-skilled networks and task forces that come together as needs dictate. Industrial R&D will have to prosper with less government subsidy and so win over those in corporate boardrooms who do not have first-hand experience of true wealth generation through innovation and high technology. □

Sir Gustav Nossal, president of the Australian Academy of Science, is in the Department of Pathology, University of Melbourne, Parkville, Victoria 3052, Australia.

Activation without a vital ingredient

David M. Chao and Richard A. Young

The complicated machinery that initiates transcription of protein-coding genes contains components, known as TAFs, that are held to be generally required for gene activation. Experiments *in vivo* change that view.

REGULATION of gene expression is fundamental to living processes, and much of it occurs at the level of transcription initiation. One form of regulation, called gene activation, produces a marked increase in the rate at which messenger RNAs are copied from the gene. The textbook model of gene activation, derived from studies *in vitro*, is that gene-specific activators bind to TBP-associated factors (TAFs), a set of proteins that associate with TATA-binding protein (TBP), thereby recruiting this component of the transcriptional machinery to promoter DNA^{1,2}. However, new evidence, reported by groups led by Michael Green and Kevin Struhl on pages 185 and 188 of this issue^{3,4}, demonstrates that gene activation can occur in living cells in the absence of functional TAFs. Does the textbook need revision?

Eukaryotic cells contain many thousands of protein-coding genes, and the task of transcribing them into messenger RNA falls to RNA polymerase II. Genes transcribed by RNA polymerase II (known as class II genes) have several promoter elements that are necessary for regulated transcription initiation. Most promoters contain a sequence element called a TATA box (because it is rich in thymine and adenine) that acts as the binding site for TBP and the rest of the transcription-initiation apparatus. The rate of transcription of each gene is regulated by activator proteins that bind to promoter DNA sequences called enhancers, generally located upstream of the TATA box. Activators, which typically have a DNA-binding domain and an activating domain, are thought to stimulate gene expression by interacting with components of the transcription-initiation apparatus. As with many other biological processes, our understanding of this is due to the combined efforts of biochemists and geneticists (see box overleaf).

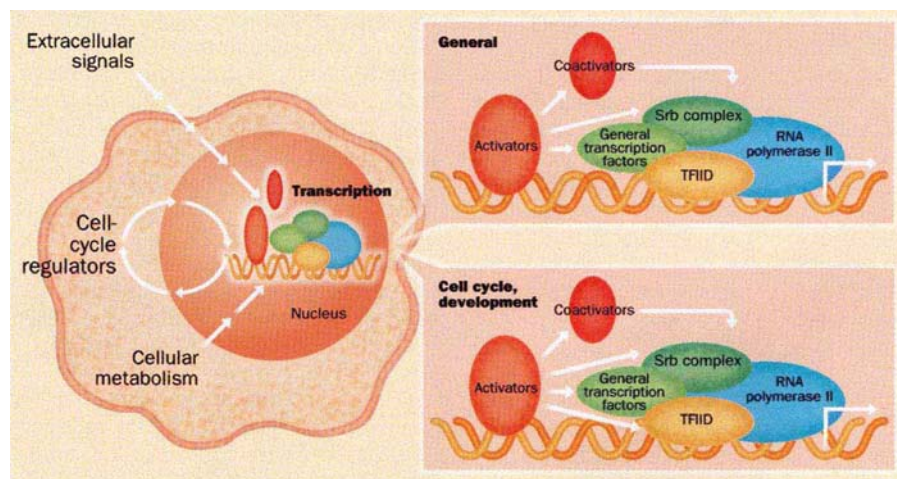
Several important features of transcriptional regulation have been re-created in the test tube with purified RNA polymerase II and a set of proteins called general transcription factors^{1,2}. These include TFIID (a complex of TBP plus TAFs), TFIIA, TFIIB, TFIIF, TFIIE and TFIIH, which are evolutionarily conserved between yeast and humans. When RNA polymerase II and general transcription factors are added to a naked DNA template, transcription can initiate at the

appropriate promoter site. The addition of an activator can, under certain conditions, increase the level of RNA synthesis.

The TAFs have been implicated in transcriptional activation *in vitro* because activators can stimulate transcription in some systems reconstituted with mammalian RNA polymerase II and the general transcription factors, but not when TBP is substituted for TFIID (refs 1, 2). The model derived from these studies is

these mutant TAFs can function appropriately at one temperature, but not another. The other approaches involved use of molecular-genetic techniques to shut down TAF synthesis and cause TAF degradation.

In each of the experiments, TAF inactivation or elimination led to cell death. Surprisingly, however, the dying cells were not compromised in gene expression generally. Moreover, activation was



Regulation of gene expression. Signals from outside and inside the cell influence the rate of transcription of class II genes. This is achieved by activators or other transcriptional regulators, which are believed to act through components of the RNA polymerase II holoenzyme, coactivators and TFIID (a complex of TATA-binding protein and associated factors known as TAFs). The involvement of TAFs is now called into question by the experiments of Green, Struhl and colleagues^{3,4}, discussed here, which show that TAFs are not generally required for gene activation *in vivo*. Rather than playing a general role, TFIID may be required for the activation of genes that participate in controlling the cell cycle and development.

that activators bind to the TAF proteins directly, thereby recruiting the TFIID complex to promoter DNA. Support for this model comes from experiments in which the response to various activators can be reconstituted with specific recombinant TBP-TAF subcomplexes. In systems reconstituted with highly purified mammalian components, cofactors such as PC4 and HMG2 are required in addition to TFIID to obtain activation *in vitro*.

The studies described by Green, Struhl and their colleagues^{3,4} were designed to find out whether yeast TAFs are generally required for class II gene expression and regulation *in vivo*. Three different experimental approaches were used to inactivate or substantially reduce the levels of TAFs in cells. In one, conditional mutations in specific TAFs were employed;

unaffected at all of the promoters tested. In contrast, mutations in most other components of the transcription apparatus have dramatically deleterious effects on class II gene expression *in vivo*. The new results raise two important questions. What factors are required for gene activation at class II promoters if the TAFs are not? And what do TAFs do *in vivo*?

Recent genetic and biochemical evidence indicates that transcription is initiated *in vivo* by a large complex of proteins called an RNA polymerase II holoenzyme^{5,6}. This apparatus has been isolated from yeast and mammalian cells, and consists of RNA polymerase II, a subset of general transcription factors and a subcomplex containing many regulatory proteins. Among these regulatory proteins are Srbs, which were genetically identified

Biochemistry and genetics: checks and balances

To study such complicated phenomena as transcriptional activation, biologists turn to two tried and true approaches — the biochemical and the genetic.

Biochemists use a strategy of reconstruction. They seek to reconstitute biological processes in test tubes. The first step is to develop an assay that reveals the activity of interest. The second is to purify the components of the cell necessary to reconstitute that activity from crude cell extracts. Finally, the purified proteins are studied individually and together to determine what each one does. Most of the three-dozen proteins required for transcription of class II genes were identified biochemically^{2,10,11}.

Geneticists, on the other hand, employ a strategy of destruction. They disrupt biological processes by making mutations and examining the consequences in living cells. The first step is to develop a screen for cells exhibiting disruption in the process being studied (transcription, pattern formation, signal transduction or whatever). The second is to pinpoint any genetic mutations associated with that disruption, then further characterize the genes concerned to determine the functions of their protein products. The fundamental mechanisms of transcriptional regulation in bacteria were worked out in this way¹².

But each approach has its limitations. Reconstituted biochemical reactions are usually a highly simplified version of what goes on inside a living cell. The reaction may be missing essential features of the process as it occurs *in vivo* (which may be poorly defined) or it may harbour hidden flaws — for example, TFIIC is missing from the roll call of general transcription factors because its identification was an artefact of the *in vitro* assay conditions.

The problem for geneticists is that the genes identified in genetic selections may not always be directly involved in the process of interest. Thus, the general transcription factor TFIIB was inadvertently identified in yeast by a screen intended to isolate translation factors. Other transcription factors have been isolated in screens for cell-cycle regulators because of their indirect effects on the expression of these regulators.

Some of these drawbacks can be overcome by sharing tactics, as we have seen in the transcription story. The chief weakness of biochemistry — inability to discriminate between what is real and what is artefact — can be addressed genetically. A biochemist with a newly purified enzyme can mutate the gene encoding that enzyme. If the biochemist is correct about the enzyme's activity, then mutation of the gene should affect the cell's

ability to perform the reaction concerned. Likewise, the chief weakness of genetics — inability to distinguish direct from indirect effects — can be tackled biochemically. A geneticist with a newly identified gene can set up assays to determine the enzymatic properties of the protein encoded by the gene. If the gene does what the geneticist thinks it does *in vivo*, then the protein encoded by that gene should have the proper enzymatic activity *in vitro*.

Of course, combining biochemistry and genetics is not always so straightforward. In some experimental animal systems, genetic manipulation techniques are not yet advanced enough to allow specific mutations to be made in a particular gene. Also, there are often difficulties in developing biochemical assays for biological processes that have not been fully defined *in vivo*.

These challenges have been met, at least in part, by molecular biologists, who use recombinant DNA technology to isolate, mutate and analyse genes and their protein products: molecular-biological tools help to bridge the gap between biochemistry and genetics. Indeed, it is now the combination of biochemistry, genetics and molecular biology that drives progress in biology, each providing checks and balances on the others. **D.M.C. & R.A.Y.**

as transcriptional regulators. Mutations in yeast *Srb* genes, unlike TAF mutations, generate global defects in class II gene expression and affect their ability to respond to regulators *in vivo*. Purified yeast RNA polymerase II holoenzyme can respond to activators *in vitro* in the absence of TAFs.

In this context, it is possible that the TAFs are indeed generally involved in the response to activators *in vivo* but are ancillary to the holoenzyme at most promoters. Activators can interact *in vitro* with components of the holoenzyme as well as TAFs^{5,6}. Contact between artificial activators and almost any component of the transcription-initiation apparatus is sufficient to activate transcription *in vivo*^{7,8}.

Another possibility is that the TAFs function at only a small number of promoters *in vivo*. Struhl and colleagues⁴ found two genes whose rate of transcription decreases when expression of some TAFs is shut down *in vivo*. It is intriguing that the promoters of both these genes have TATA boxes whose sequence deviates from the consensus. TAFs are required for transcription from such promoters *in vitro*; furthermore, at least one TAF binds directly to core promoter elements expected to be important for guiding transcription from promoters lacking consensus TATA boxes^{1,2}.

The observation that inactivation or elimination of TAFs is lethal to cells suggests that TAFs are critical for the expression of certain essential genes. TAFs appear to be involved in cell-cycle progression in yeast and mammalian cells. Green and colleagues³ point out that inactivation of particular yeast TAFs can cause growth arrest at distinct stages in the cell cycle. One of their temperature-sensitive yeast TAF_{II}145 mutants has a phenotype remarkably similar to that found in mammalian cells with a mutation in the homologue TAF_{II}250; the mammalian temperature-sensitive cell line arrests growth at the G1 stage of the cell cycle when the temperature is raised, but does not exhibit a global defect in class II gene transcription (ref. 9 and references therein).

Transcription in yeast and mammalian cells almost certainly follows the same fundamental principles, but there is no doubt that mammalian cells need to respond to more complicated and diverse regulatory signals. For example, whereas yeast have only a few developmental programmes, mammals have several hundred. If TAFs are necessary for integrating responses to diverse regulatory controls such as those governing the cell cycle and cell fate, they may be essential for expression of a higher proportion of genes in

mammalian cells than in yeast.

The final textbook entry explaining exactly how and where TAFs function *in vivo* remains to be written. Whatever the outcome, however, the history of gene-activation research underscores the importance of combining genetic and biochemical approaches. □

David M. Chao and Richard A. Young are in the Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, and the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

- Goodrich, J. A., Cutler, G. & Tjian, R. *Cell* **84**, 825–830 (1996).
- Burley, S. K. & Roeder, R. G. *Annu. Rev. Biochem.* **65**, 769–799 (1996).
- Walker, S. S., Reese, J. C., Apone, L. M. & Green, M. R. *Nature* **383**, 185–188 (1996).
- Moqtaderi, Z., Bai, Y., Poon, D., Weil, P. A. & Struhl, K. *Nature* **383**, 188–191 (1996).
- Koleske, A. J. & Young, R. A. *Trends Biochem. Sci.* **20**, 113–116 (1995).
- Halle, J.-P. & Meisterernst, M. *Trends Genet.* **12**, 161–163 (1995).
- Barberis, A. et al. *Cell* **81**, 359–368 (1995).
- Struhl, K. *Cell* **84**, 179–182 (1996).
- Wang, E. H. & Tjian, R. *Science* **263**, 811–814 (1994).
- Conaway, R. C. & Conaway, J. W. *Annu. Rev. Biochem.* **62**, 161–190 (1993).
- Zawel, L. & Reinberg, D. *Annu. Rev. Biochem.* **64**, 533–561 (1995).
- Jacob, F. & Monod, J. *J. Mol. Biol.* **3**, 318–356 (1961).

Neville Mott (1905–96)

PHYSICS is often associated in the public mind with cosmological distances, enormous particle accelerators, and abstruse theories of sub-nuclear particles: big bangs and big bucks. It is not widely enough known that the busiest speciality in physics is none of the above, but solid-state physics (now normally known as condensed-matter physics). This is the field that has brought us the transistor, the laser, modern communications — in fact, most of modern technology — and has changed our world almost beyond recognition. Starting in the 1930s, solid-state physics came about through a marriage between the infant theory of quantum mechanics and the experimental physics of ordinary materials, under the leadership of a small group of pioneers. Of these, Neville Mott, who died suddenly but quietly on 8 August in Milton Keynes Hospital, was perhaps the most influential, both through the powerful group he built up at Bristol in 1933–54, and through his many books (familarly known as 'Mott and X').

Uniquely among these pioneers, however, he moved on, first to hold positions of great academic power and influence at Cambridge (1954–73), then to open up further exciting new areas of physics, and finally to earn his Nobel Prize in 1977 (with J. H. van Vleck and me) for work done before and after his retirement in 1973 as Cavendish Professor. The field of his later work, metal–insulator transitions, became more active and controversial after the prize, rather than less; and Sir Neville waded into the thick of the action with characteristic vehemence and good humour for another dozen years or so, before spending his last half-decade on the most controversial of all fields, the theory of high-temperature superconductivity.

Neville Mott was born in 1905, of an academic family. Both his parents had been students at the Cavendish Laboratory, and his father was an educational administrator. Just too young to be in at the birth of quantum mechanics and almost entirely self-taught in the new ideas, he was in the early generation of those who exploited the theory to explain new phenomena (in his case it was relativistic scattering). But, according to his memoirs, he made a conscious decision to turn away from the increasingly esoteric directions that physics was taking, and towards practical uses of the new knowledge. Happily, at the age of 28, he was offered the professorship at Bristol, which soon would become a world centre of solid-state physics.

Notable among his activities during that period were the first attempts at understanding semiconductor physics and phenomena such as rectification, an early theory of the photographic process, and work on dislocations and the strength of materials, which began his lifelong interest in defects and disorder in solids. I well remember the spec-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

acular results on dislocations that he and Sir Charles Frank, his nephew by marriage, brought to the first great international meeting in Japan in 1953, where I first met him. (He was already chairman of the International Union at that time, and added to his impressive scientific contributions a seemingly inexhaustible store of amusing and tasteful official speechlets.)

Moving to Cambridge in 1954 to succeed Bragg as Cavendish Professor, he consolidated the initiatives in radio-astronomy and molecular biology which his predecessor had begun, built up condensed-matter theory, and added to an already strong experimental group in the field. He continued the reduction of Rutherford's speciality, nuclear physics, which he felt was much better carried on at Oxford with its close ties to the Harwell complex. Each of the fields chosen for emphasis during the two decades of his Cavendish professorship was extraordinarily fertile — Nobel Prizes were a regular occurrence.

He assumed the mastership of Gonville and Caius College, carried a full load of consulting, and was active in the Pugwash movement, along with his editorial responsibilities at Taylor and Francis. He was also an influential advisor on elementary education, especially in the former colonies.

Finally, after successfully decentraliz-

ing the administration of the Cavendish Laboratory, and giving up his Mastership after conflicts over modernizing the college, he returned to physics, concentrating on the flow of electrons in strongly interacting and disordered systems. In the 1950s he had developed the concept of the 'Mott insulator', perhaps his most important contribution of all to theoretical physics. This is a magnetic insulator where the electrons are localized on the individual atomic sites by their strong Mott–Hubbard interaction (the strong electrostatic repulsion felt by two electrons occupying orbitals in the same atom — always a stumbling block for the conventional forms of band theory).

He felt that this was not a complete enough description of insulating behaviour in the case of amorphous or impure semiconductors, however, and seized on my idea of localization by disorder scattering. To this he added several important concepts (the mobility edge, variable-range hopping, the Mott minimum metallic conductivity), but most of all he put to work what has been called his 'worldwide laboratory' of friends and collaborators, and brought the whole subject from the level of apparently esoteric speculation to general acceptance and practical use. Along the way his enthusiasm and encouragement were instrumental in bringing into being the new field of amorphous semiconducting devices such as the solar photovoltaic cell. From this spurt of interest in localization phenomena it is possible to trace a direct lineage to the now active fields of microelectronics and the quantum Hall effect.

The list of his honours is too long to relate in full, reflecting the universal respect and admiration he commanded. He was pleased to be knighted in 1962, and confided in 1977 that he was relieved not to be the first modern Cavendish Professor not to get the Nobel Prize. He was touched by his recent honorary doctorate from his own university, and in his last year was made a Companion of Honour. But no worldly honour could do justice to a scientific career that retained its fertility for 70 years, which included a major role in the creation of several entirely new sciences and a beneficent influence on several others. It touched, albeit indirectly, the lives of every one of us, scientist or not. Philip Anderson

Philip Anderson is in the Department of Physics, Princeton University, Jadwin Hall, PO Box 708, Princeton, New Jersey 08544, USA.

Fly (almost) south young bird

James L. Gould

JUST when it seemed that we understood the main principles by which migrating birds navigate, along comes a new set of experiments to make us think again. The observations, by Weindler, Wiltshko and Wiltshko (reported on page 158 of this issue¹), add a new level of complexity to our knowledge of migration, but one that actually simplifies this classic example of behavioural adaptation.

Consider the problems a young bird born in the northern temperate zone faces. It must fly to winter quarters it has never visited, leaving at the correct time, setting off in the proper direction, and often flying a dog-leg route so that it must make a critical turn to a new heading at the right time (see figure; this route skirts the Alps, Mediterranean and much of the Sahara). The flying is largely or exclusively at night, often with the ever-moving stars partially or entirely obscured by clouds, and with unpredictable winds confounding estimates of location. As the journey south progresses, the bird must deal with a shifting magnetic declination and a major change in the cast of stars visible in the sky, not to mention a slow change in their azimuth-to-time relationship as the weeks pass.

By far the most important challenge a young migrant faces is calibrating its celestial and magnetic compasses. The celestial compass provides an exact measure of true north: north is the direction of the point in the sky around which the stars appear to move over the course of the night. If a young bird memorizes the location of this pole point relative to the star patterns, then a view of a static fragment of the night sky will allow it to infer where the pole point lies, and thus determine which direction is north. Under overcast skies, the star patterns are gone, and migrants switch to magnetic cues. Unfortunately, there are few places in the breeding grounds of the higher northern latitudes (northern Europe, for instance) where true north and magnetic north lie in the same direction; the careless placement of the north magnetic pole, in upper Canada 1,400 km from the geographical pole, creates a magnetic declination, which varies with longitude and latitude.

In previous tests, migrants appeared to use celestial rotation to calibrate their magnetic compass. The migratory direc-

tion seemed to be innately programmed in terms of both cues, but when celestial and magnetic information were put into conflict (using artificial fields, or projected star patterns that rotated about a pole in the wrong direction), the birds apparently used the celestial information to recalibrate their magnetic compass². Almost inevitably, however, magnetic declination declines with decreasing latitude on a

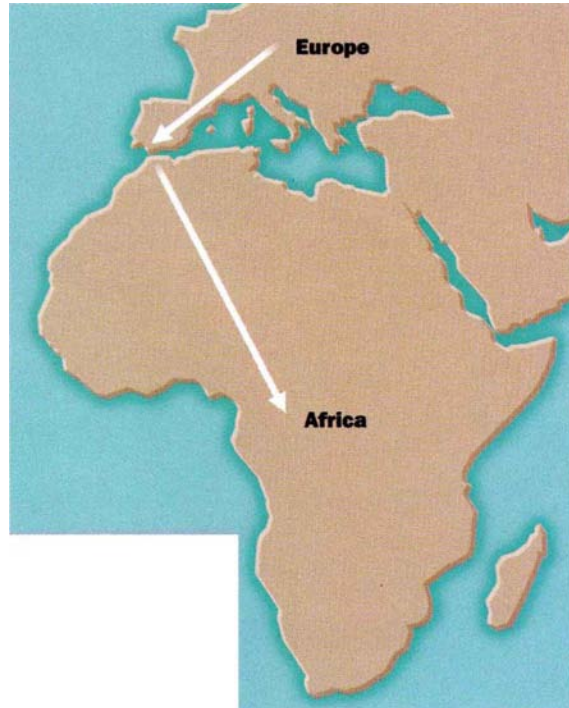


IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Dog-leg route normally followed by garden warblers migrating from northern Germany to their overwintering grounds in Africa; it avoids the Alps and the Mediterranean, and most of the Sahara. Warblers (inset) raised only with celestial (star-based) cues, as opposed to both celestial and magnetic information, instead flew due south — showing that migration direction is innately encoded in magnetic terms and, without early exposure to magnetic cues, birds fly a 'default' heading of due south.

bird's first migration, making this cue increasingly reliable. At the same time, the change in latitude and season as migration continues makes more and more of what the young bird learned irrelevant: as the bird moves south, new constellations appear over the southern horizon (as old ones vanish in the north), and new evening-sky constellations appear in the east (and disappear in the west) as the weeks of migration pass. It is no surprise, then, that recalibration *en route* is part of the picture as well³.

Weindler, Wiltshko and Wiltshko¹, in

what seems a naively simple test, compared the effect on hand-reared birds of early exposure to the normal combination of stars and magnetic direction with exposure only to the stars. They used the population of garden warblers (*Sylvia borin*) whose migratory route — south-west for several weeks, then south-east — is shown in the figure. When the time for migration arrived, both groups were tested with a static star pattern in the absence of magnetic field information. Both groups should have been able to select the appropriate direction because both had observed rotating stars during the critical period of their development. In fact, however, only the birds exposed to both celestial and magnetic information together were correctly orientated to the south-east at the outset; the birds that had grown up with just the star patterns simply flew south.

The implications of this startling result seem clear. First, celestial rotation provides only a default direction for migrants: away from the pole point (south for birds born in the Northern Hemisphere); thus, the warblers in this test reared with no magnetic information could fly due south, a generally reasonable direction but one that lacks the finesse of the normal dog-leg route. Second, the specific migratory direction — a deviation from due south — is innately coded only in magnetic terms (hence the ability of the birds reared in magnetically normal conditions to select south-east as the initial heading). Third, the magnetically encoded deviation from due south is transferred during development to the celestial compass (and so the absence of magnetic information in the tests did not interfere with the initial orientation of the birds reared with both magnetic and celestial cues, and thus with the opportunity for cue transfer). Fourth, magnetic information must be available when the time for the mid-migration turn comes; without it, even the birds that were initially well orientated lose their way.

This new picture makes sense of most previous tests, because both magnetic calibration of the celestial migration heading

1. Weindler, P., Wiltshko, R. & Wiltshko, W. *Nature* **383**, 158–160 (1996).
2. Wiltshko, R. & Wiltshko, W. *Magnetic Orientation in Animals* (Springer, Berlin, 1995).
3. Able, K. P. & Able, M. A. *Nature* **375**, 230–232 (1995).

and celestial calibration of magnetic declination normally occur. It is not yet clear whether the innately encoded magnetic deviation from a due-south migratory direction is stored in absolute or relative coordinates—that is, relative to true north (so that celestial calibration and compensation of magnetic declination must occur first), or with respect to magnetic north, so that magnetic calibration of the celestial bearing must be the initial

step. And the nature of the information essential for the mid-migration turn is an intriguing mystery. It's nice to know there are some juicy loose ends left after this major revision of our understanding of animal orientation. □

James L. Gould is in the Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544, USA.

FULLERENES

Molecular jail-making

Rolf W. Saalfrank

ONE of the most interesting features of buckminsterfullerene, or C_{60} (part *a* of the figure), is its 7-Å-diameter internal cavity, which can act as a container for atoms, and provides a potential environment for endohedral chemistry (chemistry inside the cage). Making and isolating these complexes is problematic, but a promising new pathway for gas-phase synthesis has been reported by Yves Rubin *et al.*¹

Fullerene endohedral metal complexes were first generated by Smalley and co-workers² in laser vaporization experiments on graphite doped with metal salts. Electrochemical and *ab initio* studies³ have confirmed the existence of the dimetallofullerene $La_2@C_{80}$ (that is, two lanthanum atoms trapped in a C_{80} cage). A more detailed investigation of these fascinating molecules is hindered by low reaction yields. Furthermore, the isolation of pure samples proved to be a considerable experimental challenge^{4,5}. Methods first had to be developed for separating endohedral complexes from the empty fullerenes formed at the same time. In addition, in contrast to metal-free carbon cages, endohedral metallofullerenes degrade on exposure to air.

In a very different experiment, Schwartz *et al.*⁶ showed that the noble gas atoms helium and neon can also be incorporated into C_{60}^+ and C_{70}^+ radical cations, respectively. This occurs when molecular beams of high-velocity fullerene ions are fired through a stationary noble gas. The production of neutral $He@C_{60}$ was then achieved by gas-phase reduction of

the corresponding radical cation.

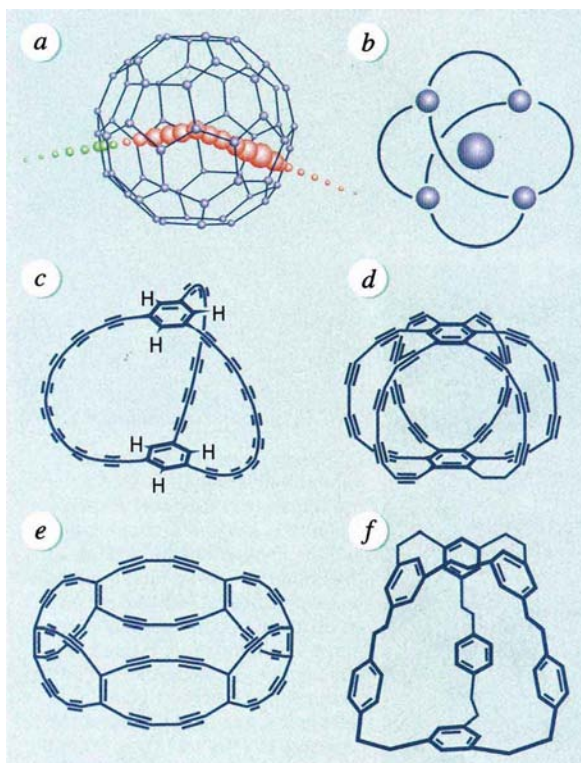
These complexes are useful because 3He has a spin of 1/2, which is excellent for NMR, so 3He NMR spectroscopy has been used⁷ to probe the interior of fullerenes. Noble gases can be introduced more rapidly into C_{60} and C_{70} by employ-

that one or more bonds in the fullerene can be broken reversibly to temporarily create 'windows' large enough for the atoms to enter.

The NMR experiments clearly show that the 3He nuclei encapsulated in C_{60} and C_{70} are electrostatically shielded by 6 and 29 parts per million, respectively, relative to free 3He . Both these numbers are unexpectedly large, and imply considerable diamagnetic ring currents inside the fullerenes—a sign of their aromaticity. These experimental findings are in agreement with *ab initio* computations^{8,9} carried out on C_{60} after the experiments. Furthermore, the calculated absolute shieldings confirm that the five-membered rings in C_{60} are paramagnetic (green spheres in part *a* of the figure) and the six-membered rings are diamagnetic (red spheres).

The first successful attachment of external groups to endohedral dimetallofullerenes¹⁰ should open the way to new organo-dimetallofullerenes. Especially interesting for future applications in materials sciences and biochemistry is the strong electron-accepting power of $La_2@C_{80}$, since its reactivity should differ significantly from that of empty fullerenes.

But substantial progress in the applicability of fullerenes and their derivatives will only be achieved if large quantities of these species are easily available. The rational synthesis of stable endohedral metallofullerenes poses a major challenge to current synthetic methods. Whereas endohedral inclusion compounds (part *b* of the figure) are readily available in supramolecular chemistry^{11,12}, that is not true for endohedral metal fullerenes. Since little more than nothing is known about the route by which fullerenes actually grow and form, it seems to be reasonable to follow the 'total synthesis strategies' that Chapman¹³ and his group have explored for buckminsterfullerene. Among them is the so-far unsuccessful preparation of the soccer ball from corannulene. The connection of three bowl-shaped corannulenes should produce a spherical $C_{60}H_{24}$ molecule, which could in



C_{60} , real and aspiring. Part *a* shows the paramagnetic (green) and diamagnetic (red) zones of the molecule, as revealed by NMR of helium caged in C_{60} . Part *b* is a cartoon of a metal-inclusion compound, with the central metal trapped by a metal-ligand cage. Similar metal- C_{60} complexes have not yet been synthesized in enough quantity to study in detail, but a few metal-trap candidates (precursors to such complexes) are shown in *c* to *f*.

ing high helium pressures. A small quantity (~100 milligrams) of a C_{60}/C_{70} mixture was heated for five hours at 600 °C in 3He at a pressure of ~2,500 atmospheres, then separated by extraction with toluene or carbon disulphide. The observed incorporation of noble gas atoms into the fullerene cavity requires

1. Rubin, Y., Parker, T. C., Khan, S. I., Holliman, C. L. & McElvany, S. W. *J. Am. Chem. Soc.* **118**, 5308–5309 (1996).
2. Smalley, R. E. *et al.* *J. Phys. Chem.* **95**, 7564–7568 (1991).
3. Suzuki, T. *et al.* *Angew. Chem.* **34**, 1094–1096 (1995).
4. Meyerhoff, M. E. *et al.* *J. Am. Chem. Soc.* **116**, 9341–9342 (1994).
5. Edelmänn, F. T. *et al.* *Angew. Chem.* **34**, 981–985 (1995).
6. Schwartz, H. *et al.* *Angew. Chem.* **31**, 605–606 (1992).
7. Saunders, M. *et al.* *Nature* **367**, 256–257 (1994).
8. Bühl, M. *et al.* *J. Am. Chem. Soc.* **116**, 6005–6006 (1994).
9. Schleyer, P. v. R. *et al.* *J. Am. Chem. Soc.* **34**, 993–995 (1995).
10. Akazaka, T. *et al.* *Angew. Chem.* **34**, 2139–2141 (1995).
11. Lehn, J.-M. *Angew. Chem.* **27**, 89–112 (1988).
12. Saalfrank, R. W. *et al.* *Angew. Chem.* **34**, 993–995 (1995).
13. Diederich, F. *et al.* *J. Am. Chem. Soc.* **31**, 1101–1123 (1992).
14. Vögtle, F. *et al.* *Angew. Chem.* **34**, 418–484 (1995).
15. Faust, R. *Angew. Chem.* **34**, 1429–1432 (1995).

theory be rearranged and dehydrogenated to produce C_{60} .

Most recently, an even more promising new strategy for the synthesis of precursors to endohedral metal fullerene complexes has been proposed by Rubin *et al.*¹ as well as by Vögtle and co-workers¹⁴. This approach is based on the remarkable propensity of a class of chemicals called cyclic polyynes to rearrange themselves into fullerenes in the gas phase. Even though the mechanism of this rearrangement is complicated, there should be several appropriate model compounds for the generation of C_{60} : 60-carbon polyacetylenic spherical cyclophanes (parts *c* and *d* in the figure), the molecular polyacetylenic carbon belt (part *e*) or the spheriphane (part *f*) are all promising candidates.

It should be emphasized that the cyclo-

phane-based hydrocarbon cages define an almost spherical cavity, which, in contrast to the inside of the fullerenes, can be accessed through custom-tailored openings on the surface of these spheriphanes. If it can be shown that the readily available corannulene is also a precursor to C_{60} , then concave cyclophane hosts with an appropriate number of carbon centres may serve as direct precursors for endohedral fullerene complexes. This new approach should have great advantages over the current strategy used to make endohedral fullerenes, using arc-discharge vaporization of doped graphite¹⁵. □

Rolf W. Saalfrank is in the Institut für Organische Chemie, Universität Erlangen-Nürnberg, Henkestr. 42, 91054 Erlangen, Germany.

QUANTUM MAGNETISM

Tunnelling secrets extracted

P. C. E. Stamp

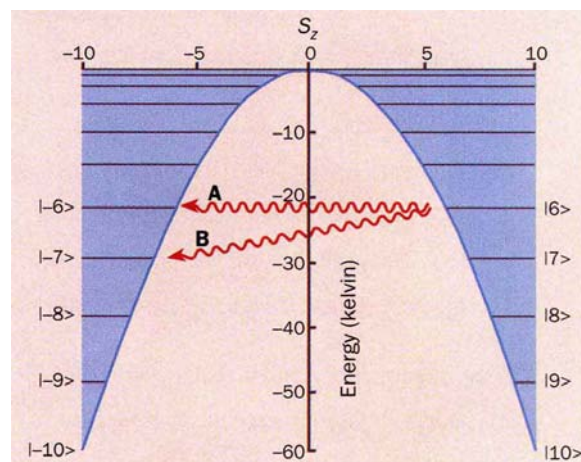
PHYSICS is, for the most part, either macroscopic and classical, or microscopic and quantum mechanical — scientists tend to shy away from the difficult middle ground. But in a remarkable synthesis of molecular chemistry and low-temperature physics, two experiments (first by Friedman *et al.*¹ and now, on page 145 of this issue², by Thomas *et al.*) have shown tunnelling between magnetization states in magnetic macromolecules arrayed in a crystal. Following earlier work by Novak and Sessoli³ and by Barbara *et al.*⁴, these experiments examine quantum effects on the treacherous frontier with classical physics, and raise interesting theoretical problems. They might also have technological implications.

The molecule in question is the manganese acetate or 'Mn₁₂-ac' molecule. One of many magnetic macromolecules that can now be synthesized⁵, this behaves at low energies like a 'giant spin', with a spin quantum number of ten. The 21 low-energy states are classified by their spin projection S_z along the crystal's symmetry axis, with $S_z = 10, 9, 8, \dots, -9, -10$, and their energy is proportional to $-S_z^2$. The energy barrier between the two lowest-energy states (with $S_z = 10$ and -10) is roughly 60 kelvin (see figure).

There must also be a term that breaks the rotational symmetry about the z axis, otherwise S_z and the energy would commute, and no transitions would be possible between the two wells (classically, angular momentum along a symmetry axis is conserved). It is a small term: at 180 millikelvin, the spin relaxation time τ of an initially polarized sample is about three years⁶. Standard theory then tells us that when the applied field is zero, transitions at low temperature between the two

extreme, low-energy states proceed through slow tunnelling at the bottom of the well; but raising the temperature should cause a sudden crossover, at a temperature of T^* , to much more rapid, thermally activated transitions between states at the top of the barrier (see figure).

A meticulous theoretical study of



The potential curve for the different spin states of the Mn₁₂-ac molecule. Orthodox theory says that at low temperatures (less than T^*) the system will remain in the two lowest-energy states, tunnelling very slowly between them; at higher temperatures most transitions occur across the top of the barrier. The new experiments show that resonant tunnelling (transitions A and B, for example) occurs between intermediate states, even at temperatures above T^* . Transition B also involves the emission of a phonon, but it still requires resonance between $S_z = 6$ and -6 .

phonon-assisted transitions in Mn₁₂-ac apparently confirms this⁷; the transitions are caused by a weak term in the energy (4th order in the spin). Assuming reasonable values for the size of this term, one easily finds that T^* is about 1.4 K. This crossover temperature is much less than

the 60 K barrier height because transitions occur much faster between levels at the top of the barrier, so even a feeble thermal occupation of levels at the top can dominate the relaxation.

But, paradoxically, the experiments refuse to obey the standard theory. Above about 1.6 K, where the system ought to be relaxing across the top of the barrier, Friedman *et al.*¹ found a remarkable series of steps in the hysteresis curve. They happen just where the applied magnetic field H (which adds $-HS_z$ to the energy, tilting the potential curve of the figure) brings into resonance pairs of levels on opposite sides of the barrier. Both groups^{1,2} find that each step is accompanied by a sharp dip in the relaxation time, falling nearly 100 times below the smooth background.

This looks like resonant tunnelling between states well below the top of the barrier, and poses a clear theoretical question — how does it happen? But that is not the only reason these experiments are important. For years, physicists have been making ever smaller magnets, searching for ones that might tunnel at low temperatures. Previous efforts were plagued by the difficulty of making ensembles of identical nanometre-scale magnets (they need to be identical because tunnelling amplitude varies exponentially with size).

Now the chemists have come to the rescue, working up from the atomic scale. The promise is of many new magnetic macromolecules, behaving like giant spins. That would truly prop open the door to the fledgling field of quantum nanomagnetism, perhaps leading to the nanomolecular engineering of magnetic memory elements that have inherently quantum properties. Such technology is of great practical interest, even if we ignore the sometimes wild speculations about quantum computers.

With this technological holy grail comes another purely academic one. Since quantum mechanics began, physicists have struggled to understand how classical physics, with its definite states, emerges from quantum physics, with its superpositions of states. Any sensible interpretation of quantum mechanics must require that at the macroscopic level of tables and chairs, classical mechanics take over — how this happens is the infamous 'measurement problem'.

Pioneering work by Caldeira and Leggett in the 1980s showed how large superconducting systems might tunnel between macroscopically different states⁸, and their predictions were confirmed experimentally

(see, for example, ref. 9). But the long-sought macroscopic quantum coherence, in which continual tunnelling between macroscopically different states is equivalent to a genuine superposition, has proved more elusive (although the biological macromolecule ferritin is a candidate¹⁰). Most physicists assume that coherent tunnelling of this kind is inevitably destroyed by interactions with the environment (phonons, photons and so on) for all but microscopic systems, leading to classical physics on macroscopic scales. The experimental confirmation of macroscopic coherence would thus force a deep re-evaluation of our most fundamental scientific theory.

The $\text{Mn}_{12}\text{-ac}$ molecule is not on the scale of tables or chairs, but it has the great advantage that, unlike large superconductors or ferritin, almost perfect samples can be prepared with identical molecules in a crystalline array, and its microscopic physics is well understood. So magnetic macromolecules, lying at the frontier between quantum physics and the 'emergent' classical physics, may lead to an understanding of how this emergence works, beyond the often vague generalities of quantum measurement theory.

So how to explain the $\text{Mn}_{12}\text{-ac}$ results? Here details are all-important; we now understand that not only phonons, but also nuclear spins, must play an essential role in the quantum relaxation of insulating nanomagnets¹¹. In $\text{Mn}_{12}\text{-ac}$, dipole interactions between the molecules may also be involved¹². The simplest explanation of these results is that the system tunnels after being thermally excited part-way up the barrier. It then appears that the nuclear spins (and perhaps also the intermolecular dipole-dipole interaction) are short-circuiting the climb to the top, even when $T > T^*$, and forcing most transitions to be via resonant tunnelling. This may resolve the paradox. Sorting out the detailed mechanism should be an interesting challenge! □

P. C. E. Stamp is in the Physics Department and the Canadian Institute for Advanced Research, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia, Canada V6T 1Z1.

Raf gets it together

C. J. Marshall

EXTRACELLULAR signals received at transmembrane receptors are relayed into the cell by signal transduction pathways. In this way signals may be greatly amplified and transmitted from the cell membrane to the nucleus, allowing appropriate responses to be elicited — for example, cell proliferation and differentiation.

One such signal transduction pathway is the Raf/Mek/ERK (extracellular-signal-regulated kinase) MAP kinase pathway, one role of which is to mediate the phosphorylation of transcription factors to change gene expression (Fig. 1). Although the activation of Mek and ERK through phosphorylation is well understood, activation of the Raf family of protein kinases (Raf-1, A-Raf and B-Raf in mammalian

phosphorylation of Raf-1 to occur, and is one mechanism of Raf-1 activation, for example by oncogenic tyrosine kinases⁶, and by X-ray irradiation⁷. However, this cannot apply to all Raf family members — B-Raf does not have the sites for tyrosine phosphorylation, and Raf-1 mutants that lack the tyrosine phosphorylation sites still have some biological activity (R. Marais and C. J. M., unpublished results).

The requirement for the Ras proteins in Raf activation is well established, but other proteins also seem to be involved. The 14-3-3 proteins bind to at least two sites on Raf-1, one at the amino terminus and one at the carboxy terminus. Microinjection of 14-3-3 into *Xenopus* oocytes leads to Raf activation⁸; and when Ras

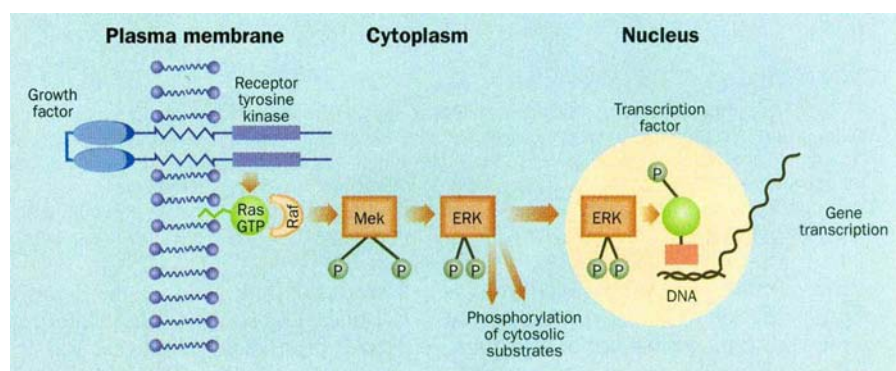


FIG. 1 The Raf/Mek/ERK mitogen-activated protein (MAP) kinase pathway, which links transmembrane receptors to intracellular signalling events.

cells) remains a puzzle. Two papers published in this issue, by Farrar *et al.*¹ (page 178) and Luo *et al.*² (page 181), now bring fresh ideas to bear on the problem.

Previous work shows that activation of Raf requires a direct interaction with the small GTPase Ras in its active GTP-bound form. But mixing Ras-GTP and Raf-1 *in vitro* does not lead to activation of Raf-1. This is in contrast to another interaction between a small GTPase and a protein kinase, where direct binding of the Rho family member Cdc42 to the p21-activated kinase (PAK) directly stimulates kinase activity, probably by inducing a conformational change that promotes autophosphorylation of PAK³. Because Ras is localized at the inner surface of the plasma membrane, the interaction of Ras-GTP and Raf within cells will bring Raf proteins to the membrane. Such a process may be a key event in Raf-1 activation, as artificially targeting Raf-1 to the plasma membrane leads to its activation in a Ras-independent manner^{4,5}.

But what takes place at the plasma membrane is not wholly clear. Recruitment to the membrane allows tyrosine

and Raf-1 are coexpressed in yeast, the activation of Raf-1 by Ras depends on the endogenous yeast 14-3-3 proteins⁹. 14-3-3 proteins are dimers, so in cells, two Raf-1 molecules could be bound to each 14-3-3 dimer. Furthermore, different members of the 14-3-3 family can heterodimerize, so heterodimers of 14-3-3 could bring about the association of Raf-1 with other proteins.

Farrar *et al.*¹ and Luo *et al.*² provide evidence that dimerization (or more likely oligomerization, because two molecules of Raf-1 are associated with each 14-3-3 dimer) may provide a mechanism for Raf activation. Both groups have induced oligomerization using the approach pioneered by Schreiber and Crabtree¹⁰. Constructs are expressed containing binding domains for bivalent drugs that cause artificial dimerization, and cells are treated with the drug. Addition of the bivalent drugs coumermycin (Farrar *et al.*) or FK1012A (Luo *et al.*) to cells expressing the Raf-1 constructs gives levels of activation equivalent to those achieved with known Raf-1 activators, for example, epidermal growth factor or phorbol esters. Farrar *et al.* also show that activation by

1. Friedman, J. R. *et al.* *Phys. Rev. Lett.* **76**, 3830–3833 (1996).
2. Thomas, L. *et al.* *Nature* **383**, 145–147 (1996).
3. Novak, M. A. & Sessoli, R. in *Quantum Tunneling of Magnetisation* (eds Gunther, L. & Barbara, B.) 171 (Kluwer, Dordrecht, 1995).
4. Barbara, B. *et al.* *J. Magn. Magn. Mater.* **140–144**, 1825 (1995).
5. Gatteschi, D. *et al.* *Science* **265**, 1054–1058 (1994).
6. Paulsen, C. & Park, J. G. in *Quantum Tunneling of Magnetisation* (eds Gunther, L. & Barbara, B.) 189–207 (Kluwer, Dordrecht, 1995).
7. Politi, P. *et al.* *Phys. Rev. Lett.* **75**, 537–540 (1995).
8. Caldeira, A. O. & Leggett, A. J. *Ann. Phys.* **149**, 374–456 (1983).
9. Clarke, J. *et al.* *Science* **239**, 992–997 (1988).
10. Awschalom, D. D. *et al.* *Science* **258**, 414–421 (1992).
11. Prokofiev, N. V. & Stamp, P. C. E. *J. Low Temp. Phys.* **104**, 143–210 (1996).
12. Burin, A. L. *et al.* *Phys. Rev. Lett.* **76**, 3040 (1996).

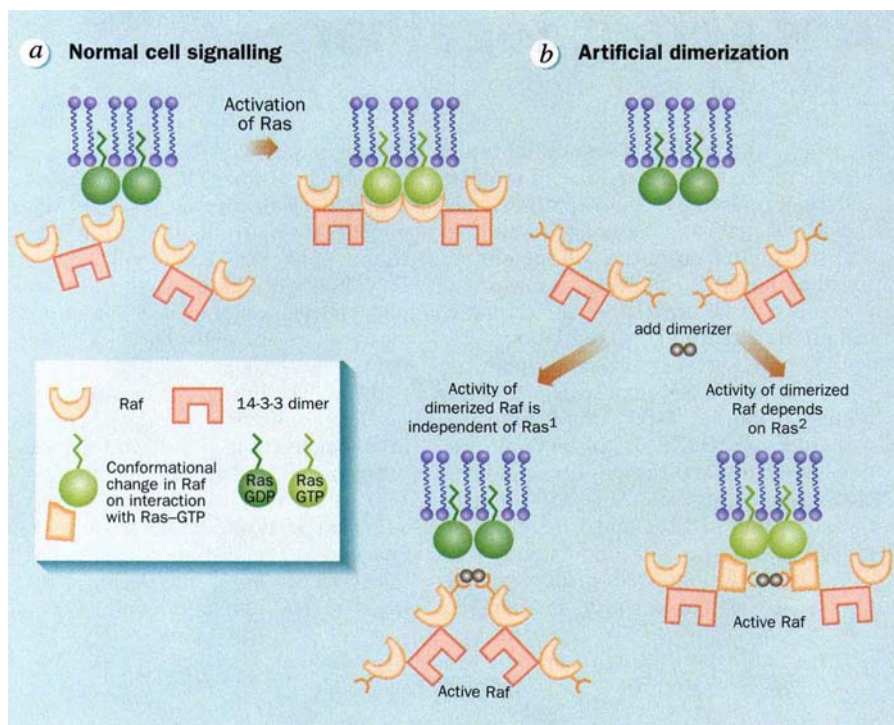


FIG. 2 *a*, Normal cell signalling. Activation of Ras to the GTP-bound form could stimulate Raf dimerization/oligomerization by concentrating Raf at the plasma membrane. *b*, Artificial dimerization. Does dimerized/oligomerized Raf still require interaction with Ras for activity? Farrar *et al.*¹ show that activation of dimerized Raf no longer requires interaction with Ras, whereas Luo *et al.*² demonstrate that dimerized Raf still requires interaction with Ras for activity. Raf is shown complexed to 14-3-3 dimers, and interaction with Ras-GTP is shown as inducing a conformational change in Raf.

oligomerization does not require tyrosine phosphorylation of Raf-1, suggesting that this is a separate mechanism of activation.

An attractive way in which oligomerization could lead to activation is through increased autophosphorylation resulting from transphosphorylation between oligomerized Raf proteins. This would provide a striking parallel with the mechanism by which receptor-tyrosine-kinase signalling occurs following growth-factor-induced receptor dimerization (see ref. 11 for a review).

The Ras-Raf interaction is critical for growth-factor-induced Raf-1 activation, so is Ras still required when Raf-1 is activated by artificial oligomerization? Here there is an interesting discrepancy between the two papers. Both groups use Raf-1 mutants that disrupt the Ras-Raf interaction, but whereas Farrar *et al.* find that the constructs that fail to interact with Ras can still be activated by chemically induced dimerization, Luo *et al.* find that their mutants do not. Even though each group has used different Raf-1 mutants, the reason for this difference is not at all clear, and this leads to alternative interpretations of the data.

Both sets of results are consistent with the idea that Ras recruits Raf to the plasma membrane, where the increased concentration of Raf at a two-dimensional surface would lead to oligomerization (see Fig. 2). However, whereas the results of

Farrar *et al.* indicate that this is all that Ras does, Luo *et al.* find that dimerization itself is insufficient to activate Raf-1; the interaction with Ras-GTP is still required — perhaps to induce a conformational change in Raf-1.

These two papers clearly demonstrate that oligomerization can lead to activation of Raf-1. But do growth factors or oncogenic Ras activate Raf-1 through this route? It has yet to be shown that growth-factor stimulation of cells leads to increased oligomerization of Raf. Nonetheless, we now have a provocative hypothesis for Raf activation that will stimulate much further work. □

C. J. Marshall is at the Chester Beatty Laboratories, Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK. e-mail: chrism@icr.ac.uk

1. Farrar, M. A., Alberola-Ila, J. & Perlmutter, R. M. *Nature* **383**, 178–181 (1996).
2. Luo, Z. *et al.* *Nature* **383**, 181–185 (1996).
3. Manser, E., Leung, T., Salihuddin, H., Zhao, S. Z. & Lim, L. *Nature* **367**, 40–46 (1994).
4. Leivers, S. J., Paterson, H. F. & Marshall, C. J. *Nature* **369**, 411–414 (1994).
5. Stokoe, D., Macdonald, S. G., Cadwallader, K., Symons, M. & Hancock, J. F. *Science* **264**, 1463–1467 (1994).
6. Marais, R., Light, Y., Paterson, H. F. & Marshall, C. J. *EMBO J.* **14**, 3136–3145 (1995).
7. Kasid, U. *et al.* *Nature* **382**, 813–816 (1996).
8. Fantl, W. J. *et al.* *Nature* **371**, 612–614 (1994).
9. Irie, K. *et al.* *Science* **265**, 1716–1719 (1994).
10. Spencer, D. M., Wandless, T. J., Schreiber, S. L. & Crabtree, G. R. *Science* **262**, 1019–1024 (1993).
11. Heldin, C. H. *Cell* **80**, 213–223 (1995).

The wired wood

SOLAR energy is hard to trap economically. Thermal systems and photocells are vastly expensive, and are useless at night. Green plants are the supreme trappers of solar energy, yet growing and burning them is primitive, inefficient and labour-intensive. Daedalus now hopes to intercept their energy more directly.

The primary stable product of photosynthesis is, of course, glucose. Ideal for plant metabolism, it is an awkward industrial fuel. But recent research has found that it can be enzymatically converted to gluconic acid and hydrogen — and hydrogen is the perfect fuel, especially for electrical fuel cells. So DREADCO biochemists are seeking to intercept plant glucose, and react it to hydrogen. They are inserting pipes into maple trees to extract the maple syrup, hydrolyse it to glucose, and react it enzymatically to gluconic acid and hydrogen. The hydrogen is bled off to a fuel cell; the gluconic acid and the non-glucose components of the sap are pumped back into the tree. The internal circulation of the sap should convey the gluconic acid to the leaves, where, with good fortune, their reductive photochemistry will intercept it, and reduce it to glucose again.

Much development will be needed to perfect the system. But the final result will be an 'electric tree', most of whose captured solar energy will appear as electricity. Nature will do all the hard work, and a buffer stock of glucose or hydrogen will keep the power flowing, even during the night. A tree subverted in this way would no longer grow, but it should still last for many decades. Whole forests of maple or other trees could be wired in this way. The small reactor and fuel cell unit hidden among the branches of each tree, and the wires looping from it, would be quite inconspicuous. Eco-freaks and landscape lovers would applaud the megawatts of natural, organic electricity flowing from apparently untouched woodland.

A tree should not be drained of all its solar power. It should be left with enough metabolic energy to grow new leaves in the spring, and to repair damage and insect attack. Daedalus even has plans to feed a bit of electricity back into the tree. His electro-osmotic scheme of last week used an electric potential to accelerate the flow of sap, especially in poor or salty soil, where roots find it hard to absorb water. His new electric trees could easily be fitted with the appropriate electrodes, and could generate their own potential. They could then flourish on useless salt marsh or contaminated land, reclaiming it as lush and verdant electric forest.

David Jones

The 1996 Antarctic ozone hole

SIR — Springtime Antarctic ozone depletion has been particularly severe since about 1992, with a loss of about two-thirds of the ozone column at the South Pole during September and near-total destruction of ozone in the 14–19-km region¹. The 1995 ozone hole was one of the longer-lasting events; however, total column ozone measurements at the South Pole for the September 19 to October 13 minimum-ozone period gave average values of 129 Dobson units (DU) compared with 109 DU in 1993 and 119 DU in 1994. The average values are precise to within a few units, so that interannual differences of 10 DU are significant. In particular, the rate at which ozone was destroyed at the South Pole in September 1995 was considerably less than in 1993 or 1994. Analysis of the trend suggests that this rate will again be high in 1996.

The rate of halogen-catalysed ozone destruction depends not only on how much halogen has been transported to the polar stratosphere, but also on the winter and spring temperatures, which can affect the amount of stratospheric cloud particle surface area available and the heterogeneous reaction rates². Interannual temperature variations, probably associated with the quasi-biennial oscillation (QBO),

in air transport from the tropics³, appear to have been present in the late 1980s and affected Antarctic minimum ozone in alternate years^{4,5} with less ozone depletion in 1986 and 1988 than in 1987 and 1989. Beginning in 1990, the QBO-related oscillation ceased, resulting in smaller variations in the amount of ozone lost from year to year in the 1990s (*a* in the figure).

There has been a general increase in the September ozone loss rate since 1986 (*b* in the figure), with periodic maxima in the 18–22-km region in 1987, 1989, 1991 and 1994. At 16–18 km, the periodicity was similar except that it was interrupted in 1992 by the enhanced depletion associated with the Pinatubo volcanic aerosol⁶, which is clearly evident below 16 km. By 1994 these aerosol effects had largely subsided.

Integrated over the 12–20-km region, the September ozone loss rate increased from about 2 to 3.5 DU per day over the past 10 years. This 75% increase in loss rate would require about a 30% increase in halogen levels if the chlorine dimer cycle² were dominant, in agreement with estimates of the stratospheric halogen increase (about 2.4 to 3.2 parts per billion from 1986 to 1995⁷). The interannual variability in the 12–20-km ozone loss rate is typically of the order of 20%, suggesting

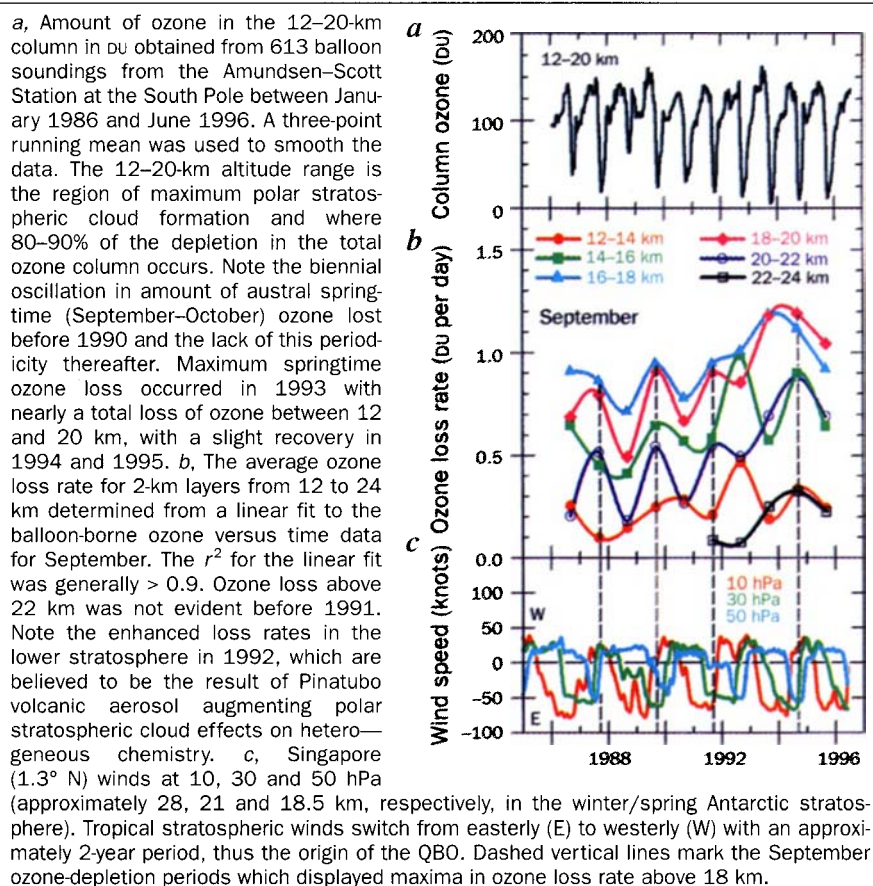
a variability in halogen levels of about 10%. The periodicity observed in the 18–22-km ozone depletion rate is matched in the QBO as indicated by tropical winds at 10, 30 and 50 hPa in the figure (*c*). No similar variations in stratospheric temperature can be detected, so the effect does not appear to be related to a modulation in polar stratospheric clouds or heterogeneous reaction rates. The high depletion rates occur during the spring following the transition of tropical winds from westerly to easterly (descending easterlies). In particular, ozone loss rates have been greater in years where the 10 hPa winds have been easterly for several months early in the calendar year preceding the ozone-hole period. Understanding the cause of this variability in ozone loss rate will be important for the detection of the expected recovery of the ozone hole early in the next century.

During the descending easterly QBO phase, transport from the tropics to the poles is enhanced⁸, and may result in the transport of more ozone-depleting halogens before formation of the winter vortex, with enhanced ozone depletion in the following spring. Enhanced winter transport at higher altitude with subsidence into the ozone-depletion region could also account for elevated ozone-depleting chemicals at this time. This might account for the more rapid September ozone decline immediately following the descending easterly QBO phase. If this hypothesis is correct, the September ozone depletion rate in 1996 should again be high because the west to east wind transition was completed in May 1996 (*c* in the figure). A steeper loss rate in September will drive ozone to lower values going into October, when the total-ozone minimum is observed. Combined with higher levels of halogens in the global stratosphere, this should result in a deeper South Pole ozone hole in 1996 than appeared in 1995, by as much as 10 DU in the late September–early October average total column ozone value.

D. J. Hofmann

NOAA Climate Monitoring and
Diagnostics Laboratory,
325 Broadway,
Boulder, Colorado 80303, USA
e-mail: dhofmann@cmdl.noaa.gov

- Hofmann, D. J. *et al. Geophys. Res. Lett.* **21**, 421–424 (1994).
- Solomon, S. *Nature* **347**, 347–354 (1990).
- Van Loon, H. & Labitzke, K. *Mon. Weath. Rev.* **115**, 357 (1987).
- Garcia, R. R. & Solomon, S. *Geophys. Res. Lett.* **14**, 848–851 (1987).
- Angell, J. K. *Geophys. Res. Lett.* **15**, 915–918 (1988).
- Hofmann, D. J. *et al. Geophys. Res. Lett.* **22**, 2493–2496 (1995).
- Scientific Assessment of Ozone Depletion: 1994, World Meteorological Organization, Global Ozone Research and Monitoring Project, Report No. 37 (Geneva, 1995).
- Trepte, C. R. & Hitchman, M. H. *Nature* **355**, 626–628 (1992).



Inferring complex phylogenies

SIR — Previous studies have demonstrated that immense molecular data sets are often needed for accurate phylogenetic estimation. For instance, under some conditions of extreme evolutionary rate variation, correct recovery of the phylogeny of just four taxa requires both an accurate DNA substitution model and information on tens of thousands to millions of nucleotides (sometimes exceeding the size of whole genomes)¹. These facts, together with the large number of possible solutions, have led many systematists to question the practicality of obtaining accurate reconstructions for phylogenies involving hundreds or thousands of species^{2,3}. Despite this scepticism, however, there is

considerable interest in phylogenetic analyses of many sequences, such as studies of the radiation of flowering plants⁴ or global diversity of human mitochondrial genomes⁵.

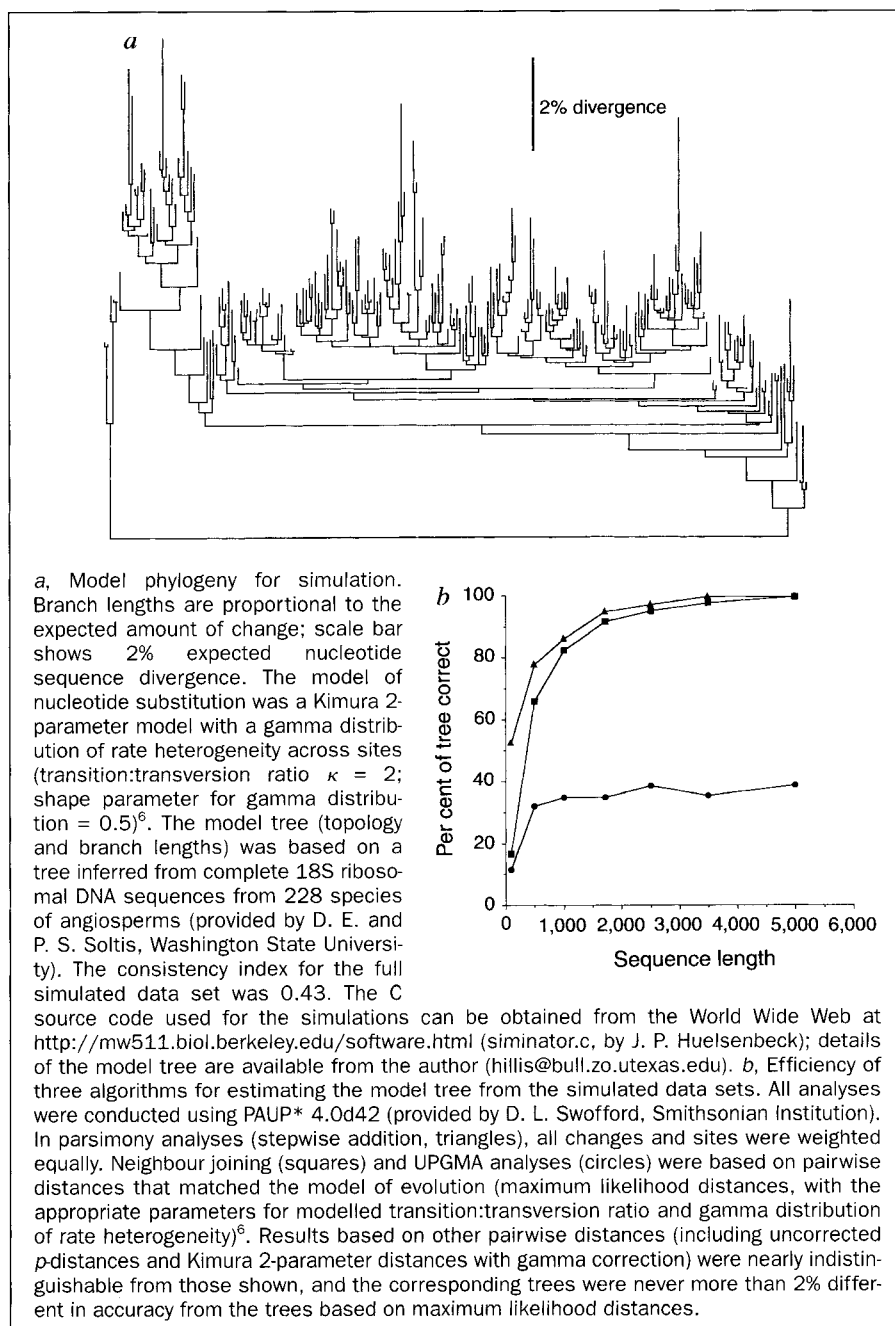
To test the feasibility of such studies and to predict the length of sequence necessary to infer complex phylogenies accurately, I simulated the phylogeny shown in part *a* of the figure. The model tree is based on an estimated phylogeny of 228 species of angiosperms inferred from complete 18S ribosomal RNA genes, and contains considerable rate heterogeneity among lineages and many short internal branches. The model of DNA evolution for the simulation included a

transition:transversion bias and rate heterogeneity across sites (see figure legend). Expected levels of change along branches were based on the observed divergences in the 18S ribosomal DNA data.

I tested three simple algorithms for approximating phylogenetic solutions: stepwise addition under the parsimony criterion; neighbour joining under the minimum evolution criterion; and the unweighted pair-group method of arithmetic averages (UPGMA) clustering algorithm⁶. The results (*b* in the figure) demonstrate that the model phylogeny can be accurately reconstructed with sequences only 5,000 nucleotides long. Considerably more nucleotides are needed to reconstruct some four-taxon trees¹.

This result is surprising because there are only three possible solutions for the four-taxon problem, but approximately 1.2×10^{502} solutions for the 228-taxon problem. Even simple methods that incorporate little of the evolutionary complexity of the model (such as parsimony) perform remarkably well (and actually outperform neighbour joining, even when pairwise distances match the model of substitution exactly). However, only the UPGMA method, which is highly sensitive to unequal rates of change⁶, failed to reconstruct more than 99% of the branches in the model tree with 5,000 nucleotides. Furthermore, branch-swapping on the stepwise-addition parsimony tree (based on the 5,000-nucleotide data set) revealed four equally good solutions, one of which matches the model tree exactly — including two unresolved polytomies (the other three solutions are various resolutions of these two polytomies). Branch-swapping on initial trees estimated from smaller data sets revealed hundreds to tens-of-thousands of equally good solutions (all of which were within 4% accuracy of the initial solution).

The explanation for the unexpected ease of reconstruction appears to lie in the dispersal of noise in the data set (homoplasy, or convergences and reversals of character states) across the many branches in the tree, so that covarying patterns of homoplasy between any two taxa are relatively rare. This allows the phylogenetic signal, even where weak, to be detected



- Hillis, D. M., Huelsenbeck, J. P. & Swofford, D. L. *Nature* **369**, 363–364 (1994).
- Patterson, C., Williams, D. M. & Humphries, C. J. *Annu. Rev. Ecol. Syst.* **24**, 153–188 (1993).
- Hillis, D. M. *Syst. Biol.* **44**, 3–16 (1995).
- Chase, M. W. et al. *Ann. Missouri Bot. Gard.* **80**, 528–580 (1993).
- Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A. C. *Science* **253**, 1503–1507 (1991).
- Swofford, D. L., Olsen, G. J., Waddell, P. J. & Hillis, D. M. in *Molecular Systematics* 2nd edn (eds Hillis, D. M., Moritz, C. & Mable, B. K.) 407–514 (Sinauer, Sunderland, Massachusetts, 1996).
- Lake, J. A. *Mol. Biol. Evol.* **4**, 167–191 (1987).
- Graur, D., Duret, L. & Guoy, M. *Nature* **379**, 333–335 (1996).
- Huelsenbeck, J. P. & Hillis, D. M. in *Molecular Zoology: Advances, Strategies, and Protocols* (eds Ferraris, J. D. & Palumbi, S. R.) 19–45 (Wiley-Liss, New York, 1996).

above the background noise. Although the amount of data needed for accurate reconstruction of such complex trees is approximately two to three times greater than has been collected to date for angiosperms, this study demonstrates that at least some problems of this order of complexity are tractable with existing methods. This simulation also demonstrates that adding large numbers of additional taxa to phylogenetic analyses may increase the accuracy of the estimated trees and at the same time reduce the need for computationally complex methods of analysis.

These results run counter to a recent trend to break phylogenetic problems into a series of four-taxon questions^{7,8}; instead,

including large numbers of taxa in an analysis may be the best way to ensure phylogenetic accuracy (although few studies to date have examined sequences long enough to provide complete resolution of complex trees). The optimal density of taxa on the tree and the necessary sequence length for complete resolution of the branching order will vary from problem to problem, so parametric bootstrapping techniques⁹ should be useful for designing full-scale studies using preliminary data sets.

David M. Hillis

Department of Zoology,
University of Texas,
Austin, Texas 78712, USA
e-mail: hillis@bull.zo.utexas.edu

Mesoderm induction in zebrafish

SIR — Mesoderm formation and its dorsoventral specification are essential processes in vertebrate early development. The mesoderm produces many types of tissues such as notochord, muscle, bone and blood cells, arranged in the correct positions along the dorsoventral axis. The mechanisms underlying these processes are best understood in the amphibian *Xenopus*, whereas little is known about this process in other vertebrates. In amphibia, the mesoderm differentiates in the equatorial region of the blastula-stage embryo through inductive interactions with vegetal blastomeres; the dorsoventral polarity of the mesoderm also depends on signals from dorsovegetal blastomeres^{1–3}. Are similar mechanisms involved in the development of teleost fish? Which tissues are

responsible for mesoderm induction in the teleosts? We have addressed these questions using zebrafish embryos.

We focused on the yolk cell because the presumptive mesoderm region in the blastula-stage embryo is formed at the vegetal marginal zone of the blastoderm, close to the yolk cell. To examine the ability of the yolk cell to induce mesoderm, we have developed a transplantation method (Fig. 1). We directly transplanted biotin-labelled yolk cells, from which the blastoderm had been removed, onto the animal-pole region of unlabelled host embryos with their yolk syncytial layer facing the host animal pole. The yolk syncytial layer, which is unique to teleosts, covers the animal region of the yolk cell under the blastoderm. During incubation, the host blastomeres covered the surface of the grafted syncytial layer. We examined the recombinants at early gastrula stages with the mesodermal markers *no tail* (*ntl*, a pan-mesodermal marker)⁴ and *goosecoid* (*gsc*, a dorsal mesodermal marker)⁵.

We transplanted yolk cells prepared from mid-blastula into the host embryos of the equivalent developmental stages. As shown in Fig. 2a, ectopic expression of *ntl* is induced as a circumferential ring around the grafted yolk cell (all positive out of the 72 operations), similar to the normal expression pattern on the host vegetal side (Fig. 2d). Histological examination reveals that ectopic expression is induced in the unlabelled host cells close to the yolk

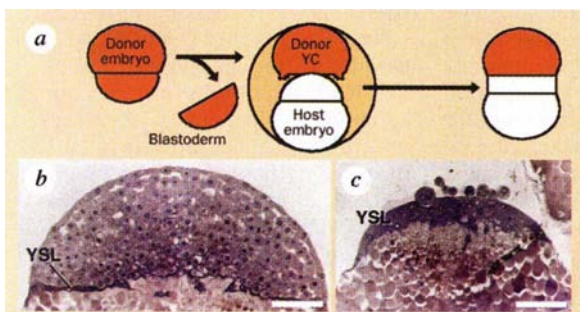


FIG. 1 Transplantation of the yolk cell (YC). *a*, Schematic representation of cell transplantation. Biotin-labelled donor cell, from which the blastoderm has been removed, was transplanted onto the animal-pole region of an unlabelled host embryo of the same developmental stage (mid-blastula stage). Recombinants were examined at 50%-epiboly with mesodermal markers by *in situ* hybridization. *b*, Light micrograph of a host embryo at the mid-blastula stage. In the marginal (MZ), yolk syncytial layer (YSL) and blastoderm cells are tightly attached to each other. *c*, Donor yolk cell. Under our experimental conditions, we could not remove the MZ cells completely with the YC intact, probably due to the tight adhesion of MZ cells with the YSL. However, those MZ cells, when transplanted into the host embryo, do not induce ectopic expression of *ntl* (data not shown). Therefore, we conclude that the few MZ cells attached to donor YCs do not affect the results obtained from our YC transplantation experiments. Scale bars, 0.1 mm.

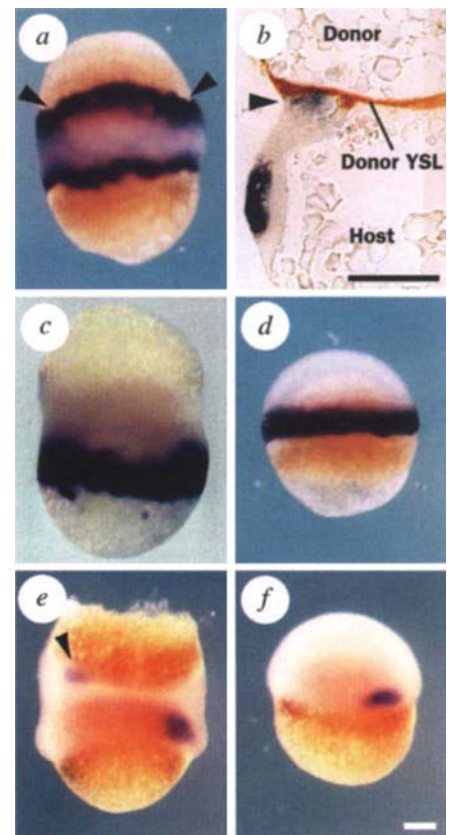


FIG. 2 Induction of *ntl* and *gsc* expression by the transplanted yolk cell. *a*, Expression of *ntl* in an embryo transplanted with a YC at the 512-cell-to-sphere stage and fixed at 50%-epiboly. Ectopic expression (indicated by arrows) is evident in a circumferential ring around the grafted YC. *b*, Histological section of a YC-transplanted embryo. Ectopic *ntl* expression (arrowhead) is seen in the host cells (unlabelled) near the donor YSL (stained brown after biotin-avidin peroxidase staining). *c*, No ectopic expression of *ntl* is observed when an unfertilized egg is transplanted instead of the YC. *d*, Expression of *ntl* in normal embryo. *e*, Expression of *gsc* in an embryo transplanted with a YC. Localized expression of *gsc* is induced near the transplanted YC (arrowhead). *f*, Normal expression pattern of *gsc* at 50%-epiboly. Scale bars, 0.1 mm.

syncytial layer of the transplanted yolk cell (Fig. 2b), suggesting that the syncytial layer is important in the induction of *ntl* expression. At all stages examined (512-cell embryo to 'sphere'), the yolk cell shows the same inducing activity. We also transplanted unfertilized eggs but found no induction of *ntl* (Fig. 2c). These results clearly show that the mesoderm-inducing signals are derived from the yolk cell and passed through the syncytial layer to the blastoderm cells.

We then examined the expression pattern of *gsc* in the recombinants. In addition to endogenous *gsc* expression, which is localized in the presumptive dorsal region, ectopic *gsc* expression is frequently induced near the transplanted yolk cell (60 out of 66 operations) (Fig. 2e, f). In most recombinants, ectopic expression is

localized in a small region that bears no relationship to the host dorsal region (statistical data not shown), indicating that dorsal tissues are induced by a localized inducing activity in the yolk cell. These data are consistent with the results obtained in other teleosts. Transplantation of younger rainbow trout blastoderms onto gastrula-stage yolk cells, from which the blastoderm had been removed, leads to development of dorsal structures by those cells located on top of the dorsal side of the host yolk cell⁶.

Our data demonstrate that, as in amphibian vegetal cells, the yolk cell of the zebrafish is responsible for induction and patterning of the mesoderm. Thus, the teleost yolk cell could be the equivalent of the vegetal endodermal cell mass in *Xenopus*. The basic mechanisms underlying the

formation of mesoderm are probably conserved between the teleosts and the amphibia. However, during the early cleavage stages in teleost development, most cytoplasmic determinants are segregated into the yolk cell, rather than the blastoderm cells⁷. For this reason, the teleost exhibits more regulative development than amphibians, keeping pluripotency until later in embryogenesis (mid-gastrula)⁸.

Toshiro Mizuno*

Etsuro Yamaha†

Masami Wakahara‡

Atsushi Kuroiwa*

Hiroyuki Takeda*§

**Department of Molecular Biology,*

School of Science,

Nagoya University,

Furo-cho, Chikusa-ku,

Nagoya 464-01, Japan

†Nanae Fish Culture Experimental Station,

Faculty of Fisheries,

Hokkaido University, 498 Sakura,

Nanae, Kameda-gun,

Hokkaido 041-11, Japan

‡Division of Biological Science,

Graduate School of Science,

Hokkaido University,

Sapporo, 060, Japan

§To whom correspondence should be addressed.

1. Nieuwkoop, P. D. *Adv. Morphogenet.* **10**, 1–39 (1973).

2. Gerhart, J. et al. *Development Suppl.* 37–51 (1989).

3. Dawid, I. B. *BioEssays* **14**, 687–691 (1992).

4. Schulte-Merker, S., Ho, R. K., Herrmann, B. G. & Nüsslein-Volhard, C. *Development* **116**, 1021–1032 (1992).

5. Stachel, S. E., Grunwald, D. J. & Myers, P. Z. *Development* **117**, 1261–1274 (1993).

6. Long, W. L. *J. Exp. Zool.* **228**, 91–97 (1983).

7. Devillers, Ch. *Adv. Morphogenet.* **1**, 397–428 (1961).

8. Ho, R. K. & Kimmel, C. B. *Science* **261**, 109–111 (1993).

Global diversity and body size

SIR — Siemann *et al.*¹ trapped an impressive variety of insects by sweep-netting the grasslands and savannahs at Cedar Creek, Minnesota. Equally impressive is their demonstration of coincident peaks for species richness (*S*) and total insect abundance (*I*) in medium-sized insects, and the strong, simple relationship ($S = I^{0.5}$) retained over a wide range of size classes. More contentious perhaps is the proposition that “if $S = I^{0.5}$ holds for nematodes, bacteria and viruses, it would suggest these phenomenally abundant small-bodied taxa might constitute most of the Earth’s diversity...”.

With respect to their first point, the discovery that most insects were of intermediate size is most easily explained in terms of functional constraints. For example, supplying body tissues with oxygen channelled through tracheae in the exoskeleton will, for physical reasons, offer optimal efficiency to insects in a particular

size range. Consider also an example from a completely different group of organisms — the discrete group of unicellular phagotrophs called ciliated protozoa. Although the variety of species occupy most of the size range possible for unicellular phagotrophs (cell lengths up to about 1 mm), most are of intermediate size — probably because they feed on the relatively narrow size range of microbial food particles and the efficiency of food particle capture depends on the predator to prey size ratio². Other functional constraints presumably explain why most birds are the size of a small thrush, and why albatrosses and hummingbirds are relatively rare and confined to specialized niches. So, within any particular group of organisms, most species lie in an optimal, intermediate size range, where there is an increased probability of dominating species appearing (for example, starlings, lemmings, krill); thus high abundance cor-

relates with high species richness when different size classes are compared.

Different groups of organisms are governed by different functional constraints, and it may be unwise to use a simple species–abundance relationship for insects for extrapolation to other groups. In the Minnesota grasslands, nematodes, mites and tardigrades may take over from insects in the smaller size range. But these groups of animals are biologically quite different from each other, and their respective ecological roles are defined by different functional constraints. It is possible that most nematode species and individuals will appear in an intermediate size class, but there is no *a priori* reason to believe the former can be predicted as the square root of the latter. The reason we believe this to be the case, is because we do now know something about a group that overlaps in size and function with the nematodes — the ciliated protozoa.

As we enter the microbial world, species richness and the abundance of individuals interact in ways that are not observed in larger organisms. Here are organisms with vast population sizes, which are readily dispersed and may be cosmopolitan, where extinctions are improbable and allopatric speciation is a rare event³. In short, we have a group in which many species may be observed in a small natural sample, although the global number of species may be modest (for free-living ciliates, probably only about 3,000 (ref. 4); for free-living marine nematodes, about 4,000 (ref. 5)), and the global number of individuals is unimaginably high.

Studies of benthic ciliate communities extending over several seasons generally reveal about 100 ciliate species (see table). This number is modest when compared with the species-rich insects from Cedar Creek; but more significant is the apparent independence of species richness from the total numbers of individuals recorded in these studies. In other words, 100 species may be close to the real total, which would change little with additional sampling effort. The pursuit of the hidden rare species only confirms that this figure is about right: fitting a log-normal distribution⁶ increases the total number of species theoretically available for detection, but only by about 10%. Moreover, when the sediment in a single core sample is manipulated in the laboratory (with various temperatures, food sources, redox gradients and so on) to provide a variety of niches that might allow rare and encysted species to grow, reproduce, and become detectable, the number of species rises, but only to 135, which appears close to the asymptote. These species included all of those recorded during the two-year sampling programme of the deep basin of the lake; so the total species richness recorded in 228 cores collected from an area of 200 m², can also be retrieved by

PATTERNS OF BENTHIC CILIATED PROTOZOA AND GRASSLAND INSECTS

Group	Location	Individuals	Species
Ciliated protozoa	Airthrey Loch, Scotland*	9,510	91
	Nivå Bay, Denmark†	17,336	91
	Esthwaite Water, England‡	20,486	104
	Helsingør Beach, Denmark‡	48,186	85
Insects	Cedar Creek, Minnesota§	89,596	1,167

*Monthly sampling over 2 years; †Sampling approximately weekly over a 1-year period; ‡Five sediment core samples removed from the deep basin every 1–4 weeks over a 2-year period; §from ref. 1.

manipulating a single core sample with an internal area of 36 cm². The species richness observed in the manipulated core represents almost 5% of all described ciliate species, and all species recorded in Airthrey Loch were also found in Esthwaite Water, which is separated by a distance of 300 km. As far as we are aware, all these species, and all those retrieved in the manipulated core, had previously been described from elsewhere in the world — supporting the widely held view that most ciliate species are indeed cosmopolitan.

Similar studies may soon be made of species richness in natural communities of bacteria and other microorganisms, and we may find that the vast number of organisms involved is indeed represented by a relatively modest global richness of species. For the moment, we await a satisfactory solution to the rather difficult question of just what is a bacterial species.

Bland J. Finlay

Genoveva F. Esteban

*Institute of Freshwater Ecology,
Windermere Laboratory,
Ambleside LA22 0LP, UK*

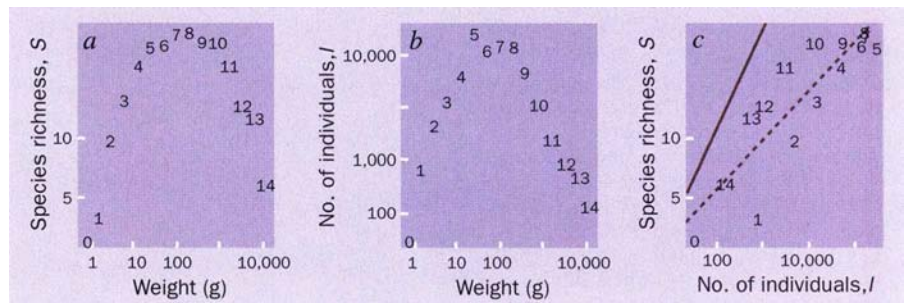
Tom Fenchel

*Marine Biological Laboratory,
University of Copenhagen,
Strandpromenaden 5,
DK-3000 Helsingør, Denmark*

1. Siemann, E., Tilman, D. & Haarstad, J. *Nature* **380**, 704–706 (1996).
2. Fenchel, T. *Ecology of Protozoa* (Science-Tech/Springer, 1987).
3. Fenchel, T. *Oikos* **68**, 375–378 (1993).
4. Finlay, B. J. *et al.* *Q. Rev. Biol.* **71**, 221–237 (1996).
5. Platt, H. M. & Warwick, R. M. *Free-Living Marine Nematodes Part 1* (Cambridge Univ. Press, 1983).
6. Preston, F. W. *Ecology* **29**, 254–283 (1948).

SIR — Siemann *et al.*¹ showed that, for 1,167 species of grassland insects, a characteristic peak in species richness (*S*) and number of individuals (*I*) occurs at intermediate body size, when species are classified into log₂ size-class intervals. The location of the peaks for *S* and *I* are approximately the same, both for the dataset as a whole, and for a given taxonomic order. Siemann *et al.* also observed a simple relationship between *S* and *I* within size classes ($S = 1.05I^{0.51}$ for the entire fauna). These observations raise the question, among others, of whether the same relationships hold for other groups². We have performed a comparative analysis for demersal fish using trawl survey data from the North Sea³. We also find that peaks in *S* and *I* are coincident (*a, b* in the figure), and observe a power-law relation similar to that found by Siemann *et al.* for the Orthoptera ($S = 1.72I^{0.25}$ compared with $S = 1.14I^{0.26}$; *c* in the figure).

Siemann *et al.* concluded that size-biased sampling is unlikely because decreases in species richness and abundance occur in all orders, even though peak body sizes differ by up to 100-fold.



a, Relation between *S* and body weight for individuals classified per log₂ weight class. *b*, Relation between *I* and body weight for individuals classified in log₂ weight classes. *c*, Relation (*S*) and (*I*) in log₂ weight classes. Dashed line, regression fit ($S = 1.72I^{0.25}$, $r^2 = 0.85$, $n = 16$, $P < 0.001$); solid line, overall fit obtained by Siemann *et al.* ($S = 1.05I^{0.51}$). (In all plots, numbers (*n*) identify individuals falling in the weight range $2^n - 2^{n+1}$ g). Data from 312 1-h trawl hauls (24 hauls per year from 1980 to 1993).

They also showed that estimates of asymptotic S_{max} (ref. 4) for within-size class rarefaction curves are close to those actually observed. We could not perform the first of these analyses because there were only 66 species of fish, so orders contain insufficient species and do not differ sufficiently in their body-size distribution. However, analyses of rarefaction curves suggest it is unlikely that we missed many smaller-bodied species.

As Siemann *et al.* note, a unimodal relation between *S* and body size is well known. What is new is that *I* should show a peak at the same location as *S* — a result common to both our analyses. This coincidence could, however, be a sampling artefact owing to the downward bias in estimates of *I* for smaller-bodied animals, resulting from mesh selectivity or differences in catchability with body size. For trawling, the mesh size we used was 30 mm when stretched open (a situation that is rarely the case when fishing), and the decline from the peak in fish abundance occurs at a body length of 70–100 mm. Mesh selectivity may, perhaps, be less likely for insect sweep-netting than fish trawling, but the possibility that individuals of small insect species are less easily dislodged from plants, or that small fish have better refuges from trawling, cannot be entirely discounted. Despite these misgivings, the patterns appear to be very similar for two widely differing taxonomic groups. Thus further efforts to determine the generality of the patterns and to seek explanations for them seem worthwhile.

There is one important difference in the way the two data sets have been treated. Siemann *et al.* assigned each individual of a given species to the body-size category of the oldest life stage, whereas we used the actual body sizes for all individuals. We used the latter approach because many younger individuals will not survive to the oldest life stage. Our data treatment does, nevertheless, present something of a paradox because, although large individuals have to be small at some stage, we do not find all species represented in

the smaller size classes. Part of the explanation for this phenomenon probably lies in the fact that many demersal fish species settle to the sea bed from a pelagic juvenile phase which would not be sampled by a demersal trawl. Because many of the species for which this occurred are relatively rare, however, it may also be that juveniles of these species occur elsewhere and only larger adults range into our sampling area.

Stephen J. Hall*

Simon P. Greenstreet

*SOAEFD Marine Laboratory,
PO Box 101,
Victoria Rd,
Aberdeen AB9 8DB, UK*

1. Siemann, E., Tilman, D. & Haarstad, J. *Nature* **380**, 704–706 (1996).
2. Nee, S. & Lawton, J. H. *Nature* **380**, 672–673 (1996).
3. Greenstreet, S. P. R. & Hall, S. J. *J. Anim. Ecol.* **65**, 577–598 (1996).
4. Colwell, R. K. & Coddington, J. A. *Phil. Trans. R. Soc. Lond. B* **345**, 101–118 (1994).

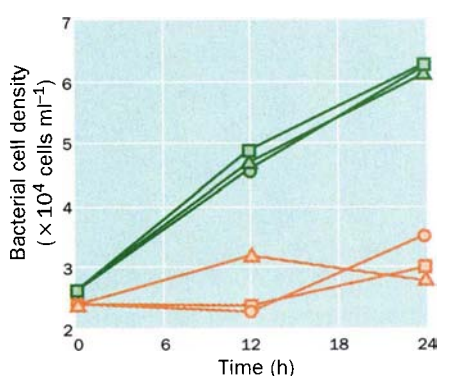
*Present address: School of Biological Sciences, Flinders University of South Australia, GPO Box 2001 Adelaide, SA 5001 Australia. e-mail: stephen.hall@flinders.edu.au

Iron stimulation of Antarctic bacteria

SIR — Recent investigations of the ocean's iron cycle have focused primarily on the response of phytoplankton to iron enrichment¹. Bacteria, however, are important in the trophodynamics and elemental cycles of marine ecosystems^{2,3}. With the exception of phototrophic prokaryotes^{4,5}, the response of bacteria to iron enrichment has largely been ignored. Here we report the results of an iron-enrichment experiment suggesting that the growth of heterotrophic bacteria in Antarctic waters is stimulated by low-concentration additions of iron.

By virtue of their abundance (10^7 to 10^9 cells per litre), high surface-to-volume ratios⁶, and ability to produce Fe-chelating siderophores⁶, heterotrophic bacteria can successfully compete with phytoplankton for available Fe. Iron stimulation of heterotrophic bacteria can lead to

FIG. 1 Growth response of Antarctic marine bacteria to Fe enrichment, 27 October 1995. Green, Fe-amended treatments; orange, unamended controls. Surface water (salinity, 34 p.p.t. (parts per thousand), temperature, -0.6°C) from the Gerlache Strait ($64^{\circ} 14.5' \text{ S}$, $61^{\circ} 47.5' \text{ W}$) was sampled with a submersible pump possessing nonmetallic internal parts and filtered through in-line $10\text{-}\mu\text{m}$ and $0.8\text{-}\mu\text{m}$ cartridge filters (Nuclepore). The plastic pump tubing and filtration equipment were flushed with copious quantities of water before sample collection. Control and Fe-amended cultures were dispensed sequentially from a single batch of water using the same HCl and sample-rinsed plastic containers. Treatments were amended with reagent-grade FeSO_4 (3.8 nM final concentration) dissolved in ultrapure water and $0.2\text{-}\mu\text{m}$ filtered before addition to cultures. Fe amendments were equivalent to an approximately 2–7-fold increase in dissolved Fe relative to ambient concentrations⁸. Controls consisted of unamended $0.8\text{-}\mu\text{m}$ filtered water. Cultures were dispensed into sterile sample-rinsed plastic tubes and incubated in the dark at -1°C , 0°C and 2°C in precision low-temperature incubators. Samples were fixed in Lugol's solution and cleared with sodium thiosulphate¹⁰ before DAPI staining and epifluorescence microscopy¹¹. Microscopic examinations indicated that only bacteria were present in cultures during experimental incubations. This work was conducted in conjunction with the DIMERS 95 research programme aboard the research vessel *Polar Duke*.



enhanced consumption of organic matter, respiratory CO_2 production and rates of bacterially mediated nutrient cycling. The response of heterotrophic bacteria to Fe enrichment, therefore, has important implications for trophodynamics and elemental cycling in marine ecosystems where Fe availability may limit biological activity.

We conducted an Fe-enrichment experiment in the Gerlache Strait during October 1995. Treatments consisted of $0.8\text{-}\mu\text{m}$ -filtered surface water amended with

nanomolar concentrations of FeSO_4 . Because temperature has a significant influence on organic substrate utilization by psychrophilic marine bacteria^{6,7}, we investigated bacterial responses to Fe enrichment at three temperatures typical of those encountered in Antarctic waters during the austral spring and summer.

Iron-amended samples exhibited an approximately twofold increase in bacterial cell densities and a fourfold increase in specific growth rate ($\mu = 0.04 \text{ h}^{-1}$) at all three temperatures compared with controls ($\mu = 0.01 \text{ h}^{-1}$) over the incubation period (Fig. 1). Because these experiments were conducted in the dark and in the absence of bacterial grazers, the data suggest that this stimulatory effect is the direct result of Fe enrichment and not an indirect effect of increased organic substrate supply from Fe-enhanced photosynthesis⁸ or the activities of bacterivores.

Iron-amended samples exhibit higher rates of leucine (Leu) incorporation per ml at near-ambient temperatures (-1 and 0°C) but not in the 2°C treatment (Fig. 2). In the -1°C treatment, cell-normalized Leu incorporation rates are similar in both Fe-amended and control samples (Fig. 2). Cell-normalized Leu incorporation rates in the 0 and 2°C Fe-amended samples, however, are lower than controls, particularly in the 2°C treatment. These data suggest that although Fe enrichment stimulates cell division (Fig. 1), Leu incorporation at higher

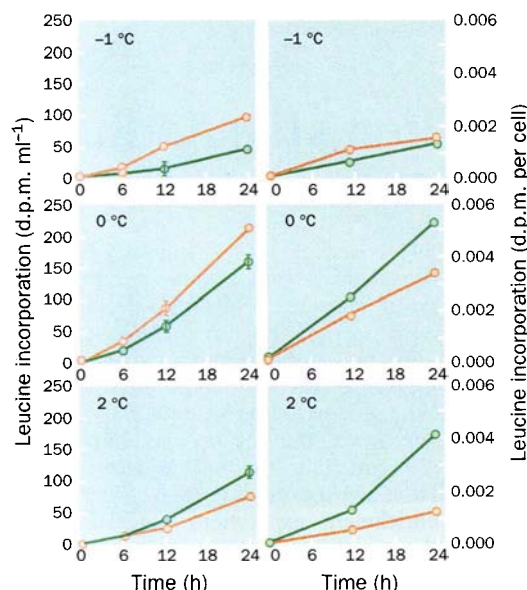


FIG. 2. Incorporation of leucine into bacterial biomass. Orange, Fe-amended cultures; green, unamended control cultures; error bars, 1 s.d. Samples were prepared as per Fig. 1 and labelled with ^{14}C -leucine (20 nM final concentration). Triplicate subsamples were obtained at each time-point, killed with 100% trichloroacetic acid and radioassayed (d.p.m. ml^{-1} , disintegrations per min per ml)¹².

than ambient temperatures in both Fe-amended and control treatments is also affected, perhaps due to temperature-induced changes in bacterial growth efficiencies or cell volumes. As we lack supporting respiration data, however, we cannot determine whether bacterial growth efficiencies are enhanced or reduced in Fe-amended samples relative to controls.

An increase in cell densities and reduction in cell-normalized Leu incorporation rates in response to Fe enrichment has previously been reported in a study⁹ where unfiltered water was amended with Fe and incubated in sunlight. Unlike the present investigation, this previous research suggested that bacterial growth is stimulated through increased supply of organic substrates from Fe-enhanced photosynthesis and not necessarily a direct response to Fe enrichment. Thus heterotrophic bacteria may respond directly and indirectly to Fe enrichment.

Our preliminary evidence suggests that the growth of heterotrophic bacteria can be stimulated by Fe additions in Antarctic waters. Our results and conclusions, however, may be relevant only to the locale and season in which these experiments were conducted. We nevertheless urge our colleagues to consider the response of heterotrophic bacteria to Fe enrichment in future ecosystem-level Fe experiments.

J. Dean Pakulski

Richard B. Coffin

Cheryl A. Kelley

Sonya L. Holder

US EPA Gulf Ecology Division,

1 Sabine Island Drive,

Gulf Breeze, Florida 32561, USA

Roswell Downer

Department of Oceanography,

Texas A & M University,

College Station, Texas 77843, USA

Peter Aas

M. Maille Lyons

Wade H. Jeffrey

Center for Environmental Diagnostics and

Bioremediation,

University of West Florida,

Pensacola, Florida 32514, USA

- Cullen, J. J. *Limnol. Oceanogr.* **40**, 1336–1343 (1995).
- Cho, B. C. & Azam, F. *Mar. Ecol. Prog. Ser.* **63**, 253–259 (1990).
- DiTullio, G. R. *et al.* *Limnol. Oceanogr.* **38**, 495–508 (1993).
- Wells, M. L. *et al.* *Limnol. Oceanogr.* **39**, 1481–1486 (1994).
- Bratbak, G. & Thingstad, T. F. *Mar. Ecol. Prog. Ser.* **25**, 23–30 (1985).
- Gottschalk, G. *Bacterial Metabolism*, 2nd edn (Springer, New York, 1990).
- Wiebe, W. J. *et al.* *Appl. Env. Microbiol.* **58**, 359–364 (1992).
- de Barr, H. J. W. *et al.* *Nature* **373**, 412–415 (1995).
- Price, N. M. *et al.* *Limnol. Oceanogr.* **39**, 520–534 (1994).
- Nishino, S. F. *Appl. Env. Microbiol.* **52**, 602–604 (1986).
- Porter, K. & Fieg, Y. *Limnol. Oceanogr.* **25**, 943–948 (1980).
- Smith, D. C. & Azam, F. *Mar. Microbial Food Webs* **6**, 107–114 (1992).

Yesterday's malaria wars

Robert S. Desowitz

The Tomorrow of Malaria. By Socrates Litsios. Pacific Press, New Zealand. Distributed by Natural History Book Service: 1996. Pp. 181. \$18 (pbk).

The Tomorrow of Malaria is, in large part, the chronicle of yesterday's malaria wars — a hundred years' conflict fought as much between malariologists as against the parasite. This is the centenary of Ronald Ross's discovery of the transmission of bird malaria through the mosquito. The super-sensitive, single-minded Ross went to his grave still holding the firm conviction that malaria could be eradicated if only weak-willed governments would commit themselves to exploit his discovery and attack the anopheline in their watery lairs.

On the other side were those who considered malaria to be a social disease. Malaria would disappear only when the economic life of the subjugated populations improved. Good housing, good nutrition, good health and educational services, and modern agricultural practices were the best antimalarials. The anopheline without the parasite was only another biting nuisance. Economic betterment was advanced as the cause of the disappearance of malaria from northern Europe and England — where more than 10,000 cases had been admitted to London's St Thomas's Hospital alone between 1860 and 1870, followed by a rapid decline to four or five cases each year by 1925.

Schismatics soon arose on each side. The technocrats divided into the chemical-therapy camp led by the German bacteriologist Robert Koch, who believed malaria should be drugged into submission, and the vector-control camp led by the Englishmen Malcolm Watson and L. W. Hackett and the Americans Fred Soper and Paul Russell, who maintained the Rossian stand of selective anopheline control. The sociocrats split into the active-betterment camp and a 'do nothing' camp which maintained, on the basis of prevailing immunological fact and theory, that holoendemic malaria led to a benign state in adults. The unfortunate wastage of the young was a necessary, worthwhile cost in the community's journey to functional immunity. Those technocrat-sociocrat wars were fought in the 'bow and arrow' era of malariology. Quinine was the only thera-

peutic drug, and pyrethrum and Paris green the only insecticides. Then the Second World War ended, the pesticide DDT was rediscovered and the World Health Organization (WHO) embarked on its global crusade to eradicate malaria.

Socrates Litsios, the WHO insider and

tion of malaria would do — impure thoughts of 'control' could lead to excommunication. A strict, universal formula for DDT application and surveillance was imposed on all national programmes. A timetable to eradication was promised.

In his book, Litsios rarely makes a direct comment or expresses an opinion. But his description of events and selection of citations speak for his sentiments. In this way, he condemns WHO for being too inflexible; for imposing a universal prescription when each setting required a programme to accommodate its vectors, its parasites and its human population's culture, behaviour and economy. Litsios faults WHO for not realizing that the massive reservoir of malaria, sub-Saharan Africa, was undoable. Nor did they appreciate the political-economic pressures on the new governments then emerging from colonial paternalism that would distract from the national malaria campaign.

The final chapter, "The Tomorrow of Today", is the sobering account of malaria's current annual toll — 300 million cases with 2 million deaths. DDT is gone. Cheap, effective chemo-therapeutics and prophylactics are gone. Multi-insecticide and multi-drug resistance are here. Steady-state, benign (for adults) holoendemic malaria has been replaced, in many settings, by unstable hyperendemicity — functional immunity impaired by the *ad hoc* chemotherapy distributed from the primary health centres. It is *déjà vu* all over again as the technocrat-sociocrat wars continue. The molecular technocrats now offer the unfulfilled, over-hyped, crime-tainted promise of a vaccine. Non-molecular technocrats offer bed nets dipped in permethrin. In one of his few displays of direct feeling, Litsios, in this final chapter, reveals himself to be the humane sociocrat: "The aim should be to improve the quality of life, with regard to human dignity, the realization of people's potentials, and the achievement of a decent standard of living. This implies a radical revision of established practices."

This is a fine, powerful little book. I would make all molecular-malaria graduate students read it before they clone their first gene or make their first hybridoma. Come to that, I would insist that their mentors read it too. □

Robert S. Desowitz is in the Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Watery lair: antimalaria programme in Binga, Zimbabwe.

sometimes apostate, has written a slim, fascinating book. Its greatest strength is its account of the WHO malaria eradication campaign beginning in 1955 and ultimately failing in 1969. Litsios recounts how the vector technocrats, now armed with DDT, took over WHO. In 1953, the psychiatrist Brock Chisholm, campaigning for re-election as director general, had remarked: "one cultural anthropologist is worth more than 100 malaria teams". The malariologists campaigned on the promise of malaria eradication. Chisholm lost and was replaced by the Brazilian malariologist Marcolino Candau. For the new WHO, nothing less than global eradica-



TAPPING lacquer from trees in late-nineteenth-century China. Taken from the latest instalment of Joseph Needham's monumental *Science and Civilisation in China*, in which C. Daniels and N. K. Menzies look at sugarcane technology and forestry. CUP, £95, \$150.

Arms control

R. V. Jones

National Military Establishments and the Advancement of Science and Technology. Edited by Paul Forman and José M. Sánchez-Ron. Kluwer: 1996. Pp. 340. £105, \$164.

FROM its title, the reader might reasonably expect to find an authoritative and broad discussion of the contributions of such establishments as those at Farnborough and Malvern in England, or at the Massachusetts Institute of Technology in the United States, or at Volkenrode in Germany, to advances in knowledge and achievement in their respective fields.

The reader might also seek enlightenment on the possible disadvantages arising from such establishments, if, for example, as happened in post-war Britain, they resulted in a dichotomy between research and teaching so that too little of the national talent was available to bring on the next generation.

The editors have not attempted a synoptic or comparative survey of the functions of research establishments but have instead invited essays from selected authors on establishments in the United States, the United Kingdom, Norway, Germany, France, Spain and Argentina.

Russia is missing from the survey, perhaps because the situation there has recently been so fluid; and an interesting, if short, story might have been told of Tito's attempt to set up a nuclear physics institute in Yugoslavia. Some of the contributions

are commendably factual; we learn, for example, that by 1890 Argentina had several institutions devoted to the naval and military aspects of science and technology.

There is a thoughtful contribution from the United States on the military origins of space sciences, which points out that much of the scientific research during the Second World War was carried out not merely by university scientists but — here was the novelty — on university campuses, and this pattern was perpetuated in the post-war years. In the United Kingdom there was no such novelty, and the government establishments remained paramount.

The German contribution, on mathematics and war, throws light on the development of aerodynamics under Prandtl and on the involvement of the mathematicians in the V-2 rocket programme and other projects based on advanced technology.

A second US contribution, both objective and highly instructive, covers quantum electronics, and in particular the development of the maser by Charles H. Townes in the Radiation Laboratory sponsored by the US government at Columbia University. This essay is full of interest, ranging from Townes's own account of how he came to invent the maser to a discussion of C. P. Snow's "euphoria of gadgets".

Snow also earns a disparaging mention in the British contribution by D. E. H. Edgerton, whose contentious approach differs sharply from the objectivity of his fellow contributors. "I will examine an important subset of British scientific intellectuals' writings: I will show that there is remarkably little correspondence between our historical picture of what they said and our historical picture of the actual relations between science, technology and war." He concludes: "There was no cumulation of knowledge. One result was that the age of the hydrogen bomb and the permanent mobilisation [*sic*] of R&D [research and development] for warlike purposes produced the vacuous, and profoundly ahistorical moralising of CP Snow and Jacob Bronowski. They influenced, unfortunately, a whole generation of 'irritating know-alls'."

Although Edgerton supports his paper with copious references, I can hardly find among them two that would have been outstandingly relevant to the discussion. The first is *The Organisation of Research Establishments* edited by Sir John Cockcroft (Cambridge University Press, 1965) and the other is *A Discussion of the Effects of The Two World Wars on the Organisation and Development of Science in the United Kingdom*, published in *Proceedings of the Royal Society A342*, 439–591 (1975).

One of the Royal Society papers deals specifically with research establishments, and it sketches the divergent views that have been advanced ever since they came into existence. The view put forward by

Alexander Strange in the 1870s, and supported by the Devonshire Commission, was that "[t]here should be established a system of national institutions for the sole purpose of advancing science by practical research quite apart from teaching it".

Lyon Playfair, although as concerned as Strange in the public fortunes of science, saw that a possible drawback resulting from state laboratories was the impoverishment of university life, because most of the research would then be done in government laboratories, and teaching would in this way be divorced from it. He pointed out that "Germany unites the function of teaching and research in the universities, whereas France keeps them in separate institutions". This latter thought might have been in the mind of Dr Phelps, the master of Sidney Sussex College, when he opposed the founding of the Cavendish Laboratory in Cambridge with the argument: "A Prussian is a Prussian and an Englishman an Englishman, and God forbid it should be otherwise".

Edgerton's 'revisionist' approach is out of harmony with the generally informative tones of the other contributions, including that of Paul Forman, who wrote the essay on quantum electronics and the genesis of the maser mentioned above, and who is also one of the two editors. Despite Edgerton's divergent viewpoint, the editors have been so impressed by his essay, which indeed has some merit, as to write in their editorial introduction:

The British scientific intellectuals thus succeeded in misleading the British public — including, Edgerton emphasizes, British historians of science and technology — even as they misled themselves. They wished to believe, and led others to believe, that, at any rate in Britain, these two enterprises, science and war, have, historically, had only an intermittent and no intimate relationship, and that, more particularly, the rate and direction of scientific and technologic progress may be discussed without close attention to society's investment in its war-making capabilities.

I am not sure about who would qualify as a British scientific intellectual, but I imagine that Francis Bacon would have been one of the first — and he wrote in 1605 that "experience doth warrant, that both in persons and in times, there hath been a meeting and concurrence in Learning and in Arms flourishing and excelling in the same men and in the same ages".

If ever I meet one of the generation of "British scientific intellectuals" as portrayed by Edgerton, I shall be tempted to follow the precept of Hermann Goering regarding 'Kultur' and reach for my gun. □

R. V. Jones, director of British scientific intelligence 1952–53, is at 8 Queens Terrace, Aberdeen AB1 1XL, UK

Humans and climate

Ian Tattersall

Children of the Ice Age. By Steven M. Stanley. Harmony: 1996. Pp. 278. \$24.95.

Paleoclimate and Evolution, with Emphasis on Human Origins. By Elisabeth S. Vrba, George H. Denton, Timothy C. Partridge and Lloyd H. Burckle. Yale University Press: 1995. Pp. 547. \$85, £60.

ARE modern humans the chance result of vagaries in climate that afflicted the African continent more than 2 million years ago? This question has long been debated by palaeoanthropologists, especially during the 20 or so years since Elisabeth Vrba, the principal editor of *Paleoclimate and Evolution*, first linked events in early human evolution to an episode of arid climate that she saw reflected in bovid fossil assemblages from early hominid sites in South Africa.

Vrba and her co-editors have now assembled a huge team of contributors, ranging from palaeoceanographers through glacial geologists to palaeoanthropologists, to examine the evidence for climate change following the early Miocene and its putative relationship to evolutionary events. The result is a mine of authoritative information, some of it new, which will certainly become a primary reference for everyone concerned with palaeoclimatology and evolution in the period addressed. Virtually all of the records that can be used as proxies of past climate are discussed, although the volume has a distinct African bias and tilts towards marine microfossils and especially terrestrial macrofossils, most notably hominids. There is also some interesting repetition, as different authors weigh more general information in relation to their pet systems.

Virtually all lines of evidence point to a stepwise deterioration in climate since the mid-Pliocene (and even earlier); and it seems fairly well established that, even at its coldest, world climate before about 2.5 million years ago (Myr) was warmer than that of today. Around that time there was a cooling event. This had a significant influence on the more cold-intolerant elements of the world's biota; so significant, indeed, that although the next major climate downturn at about 1.8 Myr was more severe in absolute terms, its biotic effects were far more muted because faunas had already been 'winterized'. Even so, the contributors to this volume are far from unanimous about the exact size, timing and effects of the ± 2.5 Myr event (for which the record is particularly spotty on the African continent); and they agree even less about whether or not it provides



MODEL of the Kissimmee River, South Florida, as used by engineers and ecologists at the University of California at Berkeley to simulate restoration of this river-floodplain ecosystem. From *Humanature*, a collection of evocative photographs reflecting our relationship with nature. University of Texas Press, \$60 (hbk), \$29.95 (pbk).

conclusive evidence for Vrba's faunal "turnover pulse" hypothesis, which she fine-tunes in this volume.

One who has no doubts whatever about the relevance of this event to human evolution is the invertebrate palaeontologist Steven Stanley, whose *Children of the Ice Age*, aimed at the general reader, is constructed around the notion that early "ape-men" were awoken from long evolutionary stagnation by the shrinking of the African forests that this episode of cooling and aridity entailed. Forced out on to the expanding open savannas, these unfortunate creatures gave "catastrophic birth" to our own genus, *Homo*. Although Stanley starts out a little smugly, he soon moves on to a smoothly flowing if perhaps overconfident account of the climate record and associated topics, and to a discussion of savanna ecology, the dangers that awaited early *Homo* in its new habitat, and the various components of what he identifies as the "adaptive complex" of this genus. As a result, his book, if questionable in the odd detail, makes a fine supplement to more conventional popular accounts of human evolution.

To discuss the origins of the genus *Homo*, of course, one has to know what *Homo* is. It has been clear for a long time that the fossils generally attributed to the early species *Homo habilis* are an oddly assorted lot, combining a variety of brain sizes and cranial and postcranial forms. Clearly, more than one hominid species is represented; yet palaeoanthropologists have been unable to agree how the fossils involved should be divided up. Both Stan-

ley and Philip Rightmire (in Vrba's tome) indicate a way out of this impasse.

Stanley rejects classical *homo habilis* from Olduvai Gorge outright, choosing *Homo rudolfensis* from Kenya's East Turkana (notably the famous cranium ER 1470, and putatively associated 'advanced' postcranial bones) as the ancestor of later hominids. Rightmire goes even further, agreeing with Ron Clarke, who has independently concluded that the fairly large-brained type specimen of *homo habilis* from Olduvai Gorge should be associated with 1470 *et al.* In that case, *Homo ergaster* from East Turkana, the first hominid definitely known to be of substantially modern body form, did not spring fully fledged from a primitive-bodied ancestor exemplified by the bulk of the 'habilis' specimens from Olduvai. Instead, it evolved (more satisfyingly) from a modestly brained hominid whose postcranial skeleton is poorly known but evidently showed advances in our own direction. The appropriate name for this species is *Homo habilis*, whose earliest representative may be a mandible from Malawi (provisionally) dated to 2.4 Myr.

This is an exciting possibility indeed, and, if correct, will inevitably re-concentrate attention on the ± 2.5 Myr event and its possible role in the origin of our own lineage. In this fast-moving field, both of these books, dissimilar as they may be, are important progress reports. □

Ian Tattersall is in the Department of Anthropology, American Museum of Natural History, New York, New York 10024, USA.

Fabian tactics

John Godfrey

Europe's Environment: The Dobris Assessment. Edited by David Stanners and Philippe Bourdeau. *European Environment Agency*; 1995. Pp. 676. £47, ECU55 (pbk). UK distributor, Earthscan.

Shaping National Responses To Climate Change: A Post-Rio Guide. Edited by Henry Lee. *Island*; 1995. Pp. 303. \$48 (hbk); \$24.95 (pbk). UK distributor, Earthscan.

Environmental Policy and Industrial Innovation: Strategies in Europe, the US and Japan. By David Wallace. *Royal Institute of International Affairs/Earthscan*. Pp. 282. £32.50 (hbk); £14.95 (pbk).

THE change in the climate that is probably now afoot is new in an important respect apart from that of being the first to be man-made. The earlier changes happened before most habitats worldwide were dominated by human action. Major changes have previously been marked by accelerated evolutionary adaptation as well as by increased extinction. We can expect the change ahead to cause more extinction than evolution.

So the systematic outline in the monumental *Europe's Environment* of environments as they now are gives a welcome baseline for one continent by which to judge what is to come. It was produced by the European Environment Agency Task Force at the behest of all of Europe's environment ministers. The large-format volume is beautifully produced, although some illustrations lack even a brief explanation.

Some chapters are masterly short surveys in their own right. That on fishing and aquaculture outlines possible changes to the European Union's Common Fisheries Policy, which has failed so far in the essential task of conserving an inexorably overexploited common resource. No prescription is given, but it is clear that a solution must be found in better control of the effort put into the catching of fish rather than, as in the past, regulating what may be officially landed. The chapter on climate change is about as good an introduction as could be encompassed in a bare ten pages. It barely mentions, however, that the European Commission has been trying to reach agreement on tax instruments to stabilize carbon dioxide emissions at 1990 levels by 2000. The proposal for a carbon tax has been stultified by industrial objections to the effect it would have on international competitiveness, and by political response to this. The regressive nature of the tax is also of political concern, exacerbated by poorer consumers spending a higher proportion of their income on fuel and likely to vote accordingly.

This theme is taken up more thoroughly in *Shaping National Responses to Climate Change*. It gives the views of a sensitive

group of US scholars on how the diverse domestic and international interests of different countries might be best accommodated to make solutions possible, rather than merely desirable. It is refreshing to read their acknowledgement that US environmental and energy policies have been particularly costly or ineffective, or both.

One chapter asks how a policy that would burden countries' economies could gain their voluntary support. It argues that the adoption of carbon taxes would not generate new vested interests that could then impede further progress, as the use of tradeable quotas would tend to do. The Organization for Economic Cooperation and Development (OECD) agrees that a tax should be the main mechanism for controlling greenhouse gases, but that a high tax applied in the countries of the OECD alone would shift carbon-intensive industries to poorer countries. Weaker regulation there might well mean that emissions actually rise.

Even if this problem can be overcome, there remains that of making an eco-tax attractive enough to be adopted while large enough to be effective. As this is not dealt with convincingly, a more subtle variant of a tax to help stabilize the climate seems worth exploring. As industry needs predictable, and at best stable, costs in order to invest rationally, it would find a tax that delivered more stable energy prices attractive. Such a tax would have to be varied in rate in inverse relation to the price of carbon fuels. In this way, the income from the tax would be greatest when the world price of these fuels was low. As the price of oil fluctuates greatly, dampening oil's oscillations would be of most help to industry. It would bring in enough revenue to reduce taxes on labour, so encouraging employment and reducing welfare costs.

A direct public benefit of an economically counter-cyclical tax would be steadier motivation to invest for a lower rate of carbon use, by any technically feasible means. A second benefit would be that if oil were to be taxed overall more than its carbon content would warrant, then oil would last longer. More of its value as a highly mobile fuel would be realized by the price it would then get. A third benefit would follow if the tax were paid into a fund to be used at a steady rate to finance the conservation and more efficient use of energy, and research and development directed towards energy projects that could be sustained.

As well as research into avoiding severe change to our climate, we also need to prepare to ameliorate its worst effects. If global warming becomes an acute problem, the market mechanism will work, because research (say into genetically engineered crops and animals adapted to harsh conditions) would pay in the short term that markets recognize. By then it will be too late for anything close to an optimal response to the problem. Even if the climate does not change as much as most now anticipate, such

research will be valuable in parts of the world already short of food, which will have even less with the increase in the world's population. This proposal for what might be called a negative-feedback carbon tax could be used to relaunch the European Union's abortive carbon tax.

Innovation of many kinds will be essential. More spending of public money is not by itself a panacea. *Environmental Policy and Industrial Innovation* uses the recent history of what has been done about key environmental problems by the richest countries to show which policies work best, and why, and which are ineffective and thus often disastrously expensive. This book will be compelling reading for anyone interested in the practical solution of environmental problems, and should be compulsory for those with responsibility.

The United States and Germany have had the most visible environmental movements. As a result they have had tough environmental policies, but these have been adversarial and inconsistent. So their policies have been comparatively costly failures because they have generated political reaction, and because of poor relations between government and industry. In the case of Germany this is odd, as on other issues government and industry relate well. By contrast, Japan and Denmark have had tough policies that have worked well, for different reasons. Japan's local governments have led the way to the development of targets that have been steadily tightened. Industry has known what was coming and prepared for it, often by innovation. In Denmark, industry has long accepted the deeply rooted environmental ethos. The Netherlands, however, is making rapid progress by combining two synergistic policies. The government has an overarching strategy for sustainable development, linked to detailed agreements between government and industry that are beginning to achieve the strategy's ends.

David Wallace leads us away from the usual slanging match between the advocates and opponents of environmental regulation. His evidence is that tough regulations that are well informed about industry lead to beneficial innovation. Unrealistic regulations are indeed a costly burden. In California, the state government forced car manufacturers by legislation to market electric cars in order to be allowed to sell the usual kind. The effect is that there are now electric second, or third, cars parked alongside the gas guzzlers.

Taken together, these books suggest possible ways to encourage international collaboration that might forestall major alteration of our climate; but they only hint at how, if this fails, richer countries may instead act to mitigate the effects of change on their own welfare. □

John Godfrey, associate of the Centre for Human Ecology, University of Edinburgh, is at 41 Lawford Road, London NW5 2LG, UK.

Circumstellar disks and the search for neighbouring planetary systems

Steven V. W. Beckwith & Anneila I. Sargent

The recent discoveries of planets orbiting several 'mature' stars bring new life to the question of just how common other planetary systems might be. Observations of very young stars provide a way to address this question and suggest that a significant number of such stars harbour conditions appropriate for the formation of planetary systems like our own.

THE discovery of other planetary systems has enormous appeal to the imagination. Inventive descriptions of other planets, almost always with intelligent life, have been the subjects of books, films, and radio and television broadcasts for many years, but they derived little from scientific research. Now the likelihood of other planetary systems, their location and their properties have become subjects of legitimate scientific inquiry; the discoveries of one planetary system¹ (the pulsar PSR1257 + 12), and candidates for extra-solar planets near several stars (notably 51 Pegasi, 70 Virginis and 47 Ursa Majoris)²⁻⁴ signify a sea-change in our ability to address the issues posed here. As we write, additional candidates for extra-solar planets are being reported. Figure 1 compares the known properties of several of these star-planet systems.

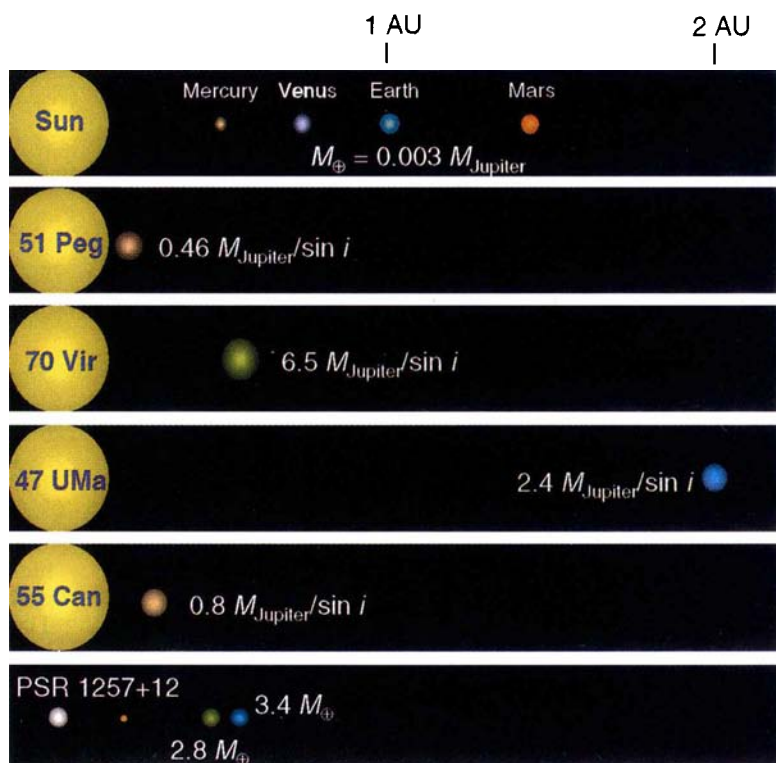
Nevertheless, the limited number of detections give few clues as to whether planetary systems are common or rare. The newly discovered 'planets' have very different properties from those in the Solar System. It is unlikely that any are hospitable to life as we know it. Direct observation of a planet is extraordinarily demanding; at present, we are limited to observing the effects of planets on a star's motion or radiation, and even these are subtle in the

extreme. Although the advent of larger telescopes and more sophisticated detection technologies should increase the number of known star-planet systems in the near future, it may be years before significant statistics can be established.

There are, however, strong hints from current research on young stars that planets could be common. Planets are thought to be assembled from a sea of small particles—dust grains—that surround a star at the time of its birth. The prevailing theory on the origin of the Solar System is that these small particles orbited the young Sun in a thin plane that extended almost continuously from the Sun's surface to well beyond the orbit of Pluto. This so-called 'disk' contained enough hydrogen gas and solids in the form of dust grains to produce the planets, moons, comets and asteroids seen today. The dust coagulated into the planets and other solid bodies during the first few million years of the Sun's lifetime and later attracted enough gas to create the giant planets like Jupiter⁵. Current theory also suggests that it is very difficult to collect all the matter from a disk into a single planet⁶; if the companions found near other stars were born from disks, they are probably members of more extensive planetary systems.

In contrast to planets, it is easy to observe disks. The total

FIG. 1 The relative separation and masses of several very low-mass companions and planetary systems that have been discovered to date. All of these objects are within the terrestrial-planet zone of a few astronomical units (AU), although most are much more massive than the terrestrial planets. The relatively small semi-major axes of the orbits mean the periods are short enough to be discovered with observations covering only a few years. Jupiter is located at about 5 AU in the Solar System, well off the right-hand boundary of this figure. Note that the companion masses depend on the unknown inclination angles of the systems (i) and are, therefore, only lower limits. ($M_{\oplus}$, mass of the Earth.) This figure is patterned after the World-Wide Web page of G. Marcy at San Francisco State University, <http://cannon.sfsu.edu/~williams/planetsearch/planetsearch.html>; this page is updated regularly with the latest candidates for planetary companions to other stars.



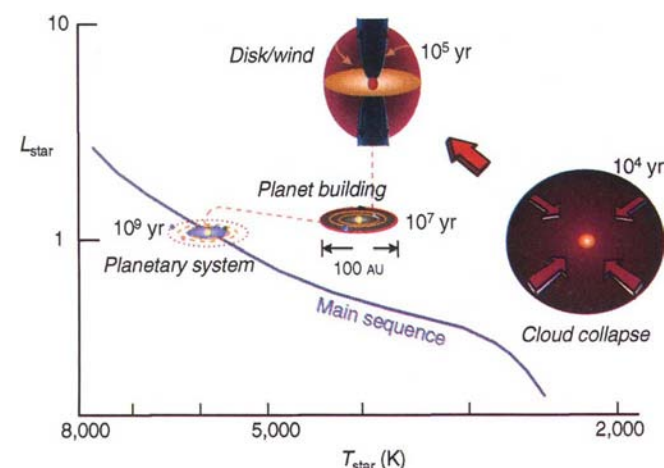


FIG. 2 The main stages of planetary formation from the collapse of an interstellar cloud (bottom right), through the active disk/outflow phase (top centre), the passive disk/planet formation phase (centre), and the existence of a stable planetary system by the time the star reaches the main sequence (left). The formation of the disk occurs naturally; it acts as a reservoir for material with too much initial angular momentum to fall onto the central star. The axes on the figure give the effective temperature and luminosity of the star as it evolves; measurements of the stellar luminosity and effective temperature are used to determine stellar age.

surface area of the small solids is of the order of 10^{14} times larger than a planet, despite their having about the same mass; disks emit and reflect light very well and can be seen to relatively large distances from the central star with modern telescopes. Starting in the late 1970s, as developments in instrumentation made their observation possible, a few stars in very early stages of evolution were seen to display features reminiscent of the primitive Solar System disk^{7–12}. It soon became apparent that disks were readily detectable and remarkably common^{13–15}. Rudimentary measures of their physical properties showed they are quite similar to our picture of the primitive solar nebula^{15,16}. This opened up the possibility of studying planetary origins, even if it was difficult to observe mature planets. If a link between disks and planets could be demonstrated, the frequency of occurrence of disks would provide an estimate of the frequency of occurrence of planetary systems.

Many of the signatures of disks are also signatures of collections of small particles near stars that may *not* reside in a thin plane. The first interpretations of observations in terms of disks depended on assumptions about plausible geometries for the matter. Although it is not always possible to check these assumptions for individual objects, there is by now so much general evidence for disks around young stars, including direct optical images, that the disk model has become well accepted^{17,18}. Also, current theories of star formation favour the star–disk configuration as an inevitable by-product of the collapse of a dense cloud of gas and dust¹⁹.

Because the conditions necessary to make planets appear to be common, we believe that planetary systems themselves are also common. In favour of this hypothesis are the many observations of disks that are close parallels to early stages of the Solar System around a large fraction of young stars. On the other hand, there is no clear proof that planetary systems will actually develop from the disks. Despite many years of searching, there are very few known objects less massive than normal stars outside the Solar System. But the disks appear to be long-lived, robust against disruption by the events that commonly accompany early stellar evolution, and very similar to our theoretical understanding of the primitive solar nebula. The few known low-mass objects may, in fact, be the tail of a much larger distribution, revealing a bias

towards finding the extreme examples with the limited capabilities of current observational techniques. If our thesis is correct, many more planetary systems should be discovered during the next few years as observing techniques improve. If some of these harbour conditions favourable to the genesis of life, what was once the stuff of science fiction will become science.

Birth of the Solar System

It is generally agreed that a flat layer of gas and dust—a disk—orbited the early Sun and provided the material which later made up the Earth, Mars, Jupiter and the other planets^{5,16,20}. The young Sun and the circumsolar disk were born from an extended cloud of gas and dust that was assembled from the detritus of dying stars and remnants of the early Universe and subsequently collapsed under its own gravity. The material accumulated quickly onto the central proto-Sun but with enough residual angular momentum to prevent some from spiralling inwards—the exact proportion remaining in orbit is not known but should have been a considerable fraction of the total mass^{17,19}. The average angular momentum of the collapsing region defined a rotation axis around which the orbits quickly stabilized, creating a disk with a thickness much smaller than its radius, at least within the regions now containing the giant planets. The formation of the stable disk probably occurred over 10^5 years after the onset of free-fall collapse²¹, almost instantaneously in cosmic time.

As the central proto-Sun evolved, the solid particles in the orbits settled to a dense layer in the mid-plane of the disk and began to stick together as they collided^{15,22,23}. During the next 10^4 – 10^5 years, large rocks and small asteroids grew gradually from the small dust particles²⁴. When the gravitational pull of the largest asteroids was sufficient to attract neighbouring pebbles and rocks, they grew even more rapidly to the size of small planets²⁵. The terrestrial planets are large accumulations of solid particles that grew from the collisions between these planetesimals. In the outer parts of the disk, a few such solid cores became large enough (~ 10 Earth masses) to accrete gas^{26,27}, the dominant reservoir of mass, and gave rise to the giant gas planets²⁸. Temperatures close to the proto-Sun were presumably too high to allow gas accretion.

The planet-building phase is thought to have taken between $\sim 10^7$ and a few times 10^8 years. The inner planets probably grew quickly, whereas the most distant gas giants required a very long time; current theory still has some difficulty explaining the existence of Neptune. The genesis of life on Earth happened within a billion years after this phase²⁹. A schematic outline of Solar System development is shown in Fig. 2.

Disks and planetary systems

What characteristics of disks are prerequisites for the creation of planetary systems? By analogy to the Solar System, we expect a planet-forming disk should have more than 0.01 solar masses ($0.01 M_{\odot}$) of gas distributed within approximately 50 astronomical units (AU) of a young star. (1 AU is the distance between the Sun and the Earth.) It should be free from disruption for at least a few million years to create rocky cores, and perhaps longer to build the gas giants. But these characteristics probably represent only a subset of the disks that accompany star formations. Conceivably, a disk with substantially lower mass ($10^{-6} M_{\odot}$), size (a few AU) and lifetime ($\sim 10^6$ years) could create terrestrial-type planets suitable for life without the presence of gas giants. In principle, even larger, more massive disks could accompany the birth of stars much more massive than the Sun.

The young disks we describe here have masses, sizes and lifetimes that allow the build-up of planets from dust; they are, in these respects, very similar to the disk around the young Sun. The inner regions in many young disks appear to be cleared early in their evolution and remain cleared, as would result from the gravity of larger bodies acting to perturb the orbits of the small particles. They are presumably the progenitors of the older disks seen around mature stars such as Vega and β Pictoris³⁰. The older disks have not dissipated completely, as might be expected if some

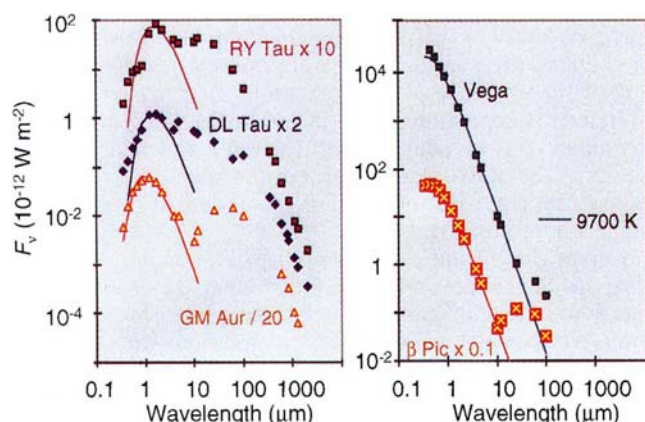


FIG. 3 Spectral energy distributions of the young stars RY Tauri, DL Tauri and GM Aurigae (left) and the main-sequence stars Vega and β Pictoris (right). The flux density, F_v , is a density in frequency units, ν . The solid lines in each figure are Planck functions at the effective temperatures of the stars: they represent emission from the stellar photospheres. The data points show thermal infrared emission from the small particles in the disks. The ratio of disk to stellar emission decreases among this sample as: RY Tau, DL Tau, GM Aur, β Pic, Vega. The relatively diminished disk emission at wavelengths less than $\sim 20 \mu\text{m}$ in GM Aur is interpreted to mean that there are no solid particles within $\sim 1 \text{ AU}$ of the star; a central hole has developed.

disruptive event blew all the matter away from the stars. To maintain the supply of short-lived smaller particles in the older disks, there is likely to be a reservoir of bodies colliding with each other as in the accumulation stage of planet-building. There is good circumstantial evidence for planet building in young and old disks alike, and disks are a common occurrence around young stars.

Observation of the disks

It is relatively easy to find young disks by observing their strong emission at infrared wavelengths. The infrared radiation is emitted by small solid particles within the disks and is several orders of magnitude larger than the emission from the stars³¹; Fig. 3 shows several examples of the emission from star-disk systems at different stages of their evolution. Although the solid particles probably make up no more than about 1% of the disk mass—the rest being gas-phase hydrogen and helium—they produce almost 100% of the radiation, owing to their continuous emission between about $2 \mu\text{m}$ and $\sim 1 \text{ mm}$ in wavelength. The range of

strong emission from the near infrared to the radio is the result of a wide variation of temperatures, from $\sim 1,000 \text{ K}$ very close to the stars to $\sim 30 \text{ K}$ near the outer edge of the disks³¹ several hundred AU away.

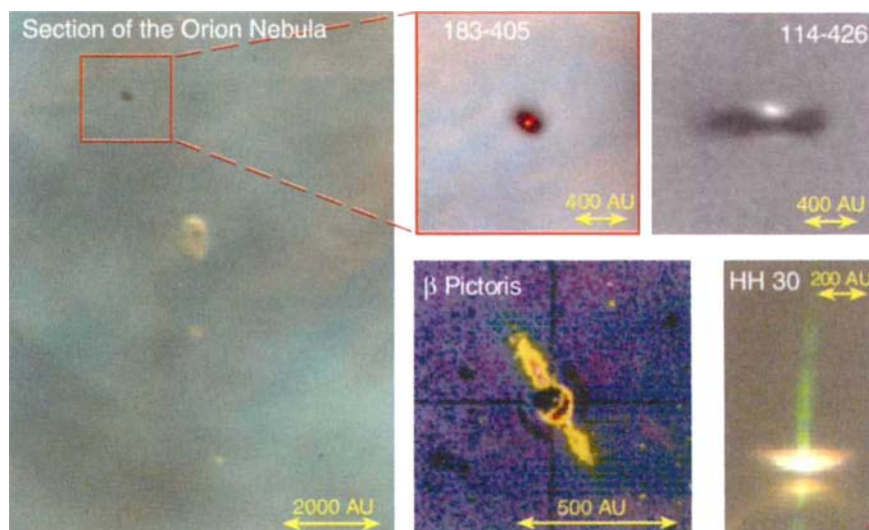
Detection of infrared emission is, nevertheless, an indirect way of finding disks and not by itself an unambiguous demonstration that the solid particles are in orbit around the stars. Fortunately, it has recently become possible to get direct images of disks with the Hubble Space Telescope; several examples are shown in Fig. 4. A number of disks can be seen in silhouette against the bright background of the Orion nebula³². These silhouettes show flattened, elliptical shapes and demonstrate for the first time that the disks are truncated sharply at their outer radii³³. The bulk of the mass is confined to a small area, but the measured disk sizes are several times larger, because the opacity of dust particles is large at optical wavelengths; the silhouettes include rather tenuous material in the most distant parts of the disks. Occasionally, disks are viewed nearly edge-on, and the scattered light illuminates the flaring outer structure³⁴, with HH 30 in Fig. 4.

The images presented in Fig. 4 are striking confirmations of the disk geometries inferred from observations at infrared and millimetre wavelengths. Whereas there was always some doubt about the distribution of material based on spectral energy distributions such as those shown in Fig. 3, the optical images make a compelling case for the basic disk geometry: thin planes in the inner regions ($< 50 \text{ AU}$) flare gently at several hundred AU and are often truncated sharply at their outer edges.

In fact, most disks bear a striking resemblance to the standard model of the primitive solar nebula described in the last section. More than 100 have now been observed at millimetre wavelengths where the thermal continuum radiation penetrates through the mid-planes and allows an assessment of the total amount of solid matter in orbit^{15,35,36}. The average total mass in the solid particles, containing almost all elements heavier than helium, is between 10^{-4} to $10^{-3} M_{\odot}$, essentially the same as inferred for the primitive solar disk and at least ten times that needed to make a planetary system like ours—the total mass in elements heavier than helium in the Solar System is of the order of $10^{-4} M_{\odot}$. Assuming 100 times as much hydrogen gas associated with the particles (the characteristic ratio for the interstellar medium), typical disk masses are between 0.01 and $0.1 M_{\odot}$, the range normally attributed to the primitive solar nebula. A distribution of the masses inferred from the particles is shown in Fig. 5. The typical disk mass is equal to that inferred for the disk around the young Sun before the formation of the planets.

Attempts to measure the gas masses from molecular lines of trace molecules, such as CO, indicate gas masses at least ten times smaller than those derived from the particles³⁷⁻⁴¹. The uncertain-

FIG. 4 The left-hand image is a small section of the Orion nebula taken with the Hubble Space Telescope by O'Dell and Wen³²; the box encloses a disk seen in silhouette, which is expanded (top-centre image) as object 183-405 (HST16) observed by McCaughrean and O'Dell³³. Top-right image is another, similar disk seen almost edge-on; the faint light above and below the disk plane is scattered from particles above and below the disk. Notice that whereas the central star is visible in 183-405 (disk is nearly face-on), it is obscured from view in 114-426 (top right) and HH30 (bottom right) which are both nearly edge-on disks. Bottom centre, image of scattered light from the disk around β Pictoris taken by Smith and Terrile⁹. The bottom-right figure is another Hubble Space Telescope image of a bowl-shaped disk near HL Tau seen almost edge-on by Burrows *et al.*³⁴; the green jet emerging from the centre is part of an outflowing wind (compare Fig. 2), and the light above and below the disk plane is scattered from tenuous particles probably falling onto the disk.



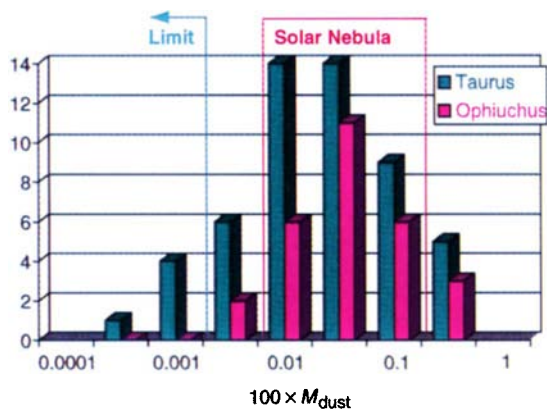


FIG. 5 Distribution of disk gas masses $100 \times M_{\text{dust}}$ (in solar masses) for two samples of stars in the dark clouds in Taurus^{15,35} and Ophiuchus³⁶. The red box denotes the range of estimates for the mass of the primitive solar nebula before the planets were created. 'Limit' refers to the smallest mass than can be reliably observed.

ties involved in calculating gas masses are large, especially as the emission is often detected from only the disk surface and not the entire volume, and many of the trace molecules than can be observed may be frozen onto the particles. But there is some indication that the gas is depleted relative to the dust in disks, thus making it difficult to create giant planets such as Jupiter.

Studies of large samples make it clear that disks surround many young stars^{13–15,35,36}. Approximately half of the classical T Tauri-type stars in the Taurus-Auriga and Ophiuchus dark clouds show evidence for disks. The presence of weak-line T Tauri stars— young stars without disks⁴²—in the same dark clouds could lower this percentage to between 10 and 30%, but studies of proper motions (the apparent change of position with time) cast doubt on whether these additional stars are *bona fide* group members⁴³. A most important recent finding is that disks are common in dense clusters of young stars^{44–46}, such as at the centre of the Orion nebula⁴⁷, where more than 1,000 stars exist within a region a few light years across—that is, the distance between the Sun and the next nearest star—and where the nebular gas is a plasma at 10,000 K. Most stars in the Galaxy are born in dense clusters, and most stars are members of binary or multiple star systems even after they move away from the cluster environment. There is evidence for disks in 70–80% of the young stars in these dense clusters⁴⁷ and also in many isolated binary systems. This evidence could be made into convincing proof with very high resolution images of some of these cluster star disks.

The bulk of the matter in disks is within a few tens to a few hundreds of AU from the stars with the majority distributed over ~ 100 AU. Observations of close binary stars^{35,48}, analyses of the spectral energy distributions^{15,31} and observations that resolve the thermal emission at millimetre wavelengths^{49,50} show that the major concentration of material is within the equivalent size of Pluto's orbit (80 AU diameter). Tenuous material can often be traced to much larger radii. Emission lines of molecules in the gas phase have been seen to several thousand AU from some stars^{37,38,51}, and the scattered light from the disk around β Pic-toris^{9,52} extends to more than 2,000 AU.

Many of the particles in the young disks survive for at least a few million years^{15,33}, more than ten times the time for build-up of small particles into larger rocks and asteroids in the young Solar System. There is, however, an absence of near-infrared (2–10 μm) emission from some disks that have considerable emission at longer wavelengths indicating material at distances greater than 1 AU. This lack of emission is normally attributed to holes that develop within 1 AU of the star, where the disk material is either absent or very tenuous^{53,54}. There have even been interpretations requiring gaps⁵⁵ in the matter distribution, as were thought to

occur in primitive solar-nebula models. Whereas gaps can equally well be explained by sharp temperature jumps in the disk owing to the condensation of ices⁵⁶, the inner holes that would develop naturally as large bodies are built up in orbit cannot be easily attributed to temperature effects in a continuous disk. Inner holes larger than 10 AU in radius may be resolved in the main-sequence phase^{30,52}. The much smaller holes in the young disks might be resolved directly with the coming generation of ground-based optical interferometers.

The particles in some young disks have different properties than the grains in the interstellar medium: either they are larger on average or have undergone some chemical processing. The result of this processing is a change in the emissivity—the capacity to emit light at a specific wavelength—at millimetre wavelengths. For a fraction of the known disks, the emissivity is constant or decreases slowly with increasing wavelength at wavelengths longer than 1 mm (refs 57, 58). Laboratory measurements^{59,60} and calculations⁶¹ of the intrinsic opacity variations in typical interstellar and interplanetary materials show that although the opacity decrease need not be as rapid as λ^{-2} , where λ is the wavelength, it is never less rapid than λ^{-1} . Quite a few disks are seen with opacity decreases proportional to $\lambda^{-\beta}$ with $\beta < 1$. It is not yet known, especially for $\beta = 0$, whether the opacity itself is also relatively small—that is, the disks are grey (partial absorption) or black (total absorption) at these long wavelengths; black absorption

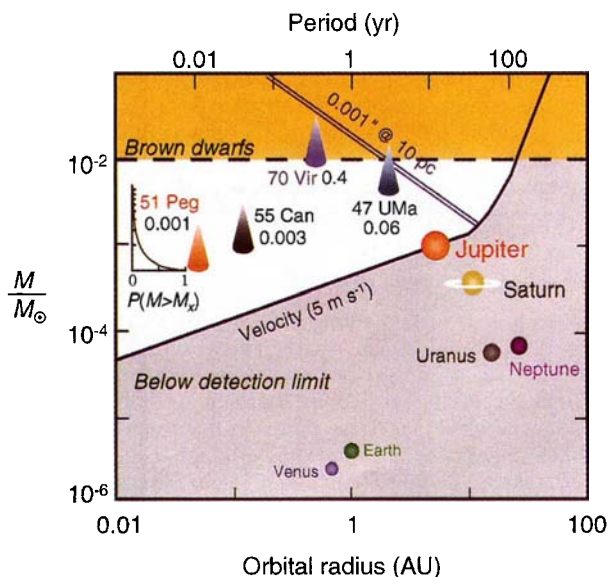


FIG. 6 Comparison of the masses and orbital semi-major axes of low-mass stellar companions. The tan region at the top shows the mass range for brown dwarfs. The masses of the companions to 51 Peg, 47 UMa, 70 Vir and 55 Can are lower limits, due to the uncertainty in orbital inclination. The probability that the actual mass is greater than a certain value is plotted next to 51 Peg and applies to all three newly discovered stellar companions. The number next to the object label is the *a priori* probability that the companion mass is greater than $0.01 M_{\odot}$, that is, it is in the brown-dwarf range. The symbols are shaded in brightness to indicate the relative probability of the actual masses, with darker meaning more probable. The blue double lines show the current sensitivity of astrometric measurements to detecting companions of solar-mass stars 10 pc away (the observable region is above and to the right of the line)—1 milliarcsec is currently the best accuracy, but 10 microarcsec will be possible with the largest telescopes and optical interferometers. The dark blue line shows the sensitivity of radial velocity observations (the observable region is above and to the left of the line)—5 m s^{-1} is the current state of the art. The grey region is not accessible by current observational techniques. Planets more distant than ~ 10 AU from the stars probably would not have been discovered by now, owing to their long orbital periods; this insensitivity is indicated by the upturn in the velocity line. Jupiter-like planets are almost detectable around other stars, but it is not yet feasible to search for Earth-like planets using these techniques.

would occur for a large density of small particles. Observations with millimetre-wave interferometers are beginning to address this issue^{49,50}.

There is some evidence that the particles surrounding the main-sequence stars are larger on average than typical interstellar grains⁶² but otherwise of similar composition⁶³. These disks have, on average, several orders of magnitude less mass in small particles than their younger counterparts. The disk masses around the main-sequence stars are typically $10^{-7} M_{\odot}$, although β Pic at $\sim 10^{-3} M_{\odot}$ is several orders of magnitude more massive than the average. Between 20 and 30% of the main-sequence stars probably have tenuous disks³⁰. The particle lifetimes in these older disks are much shorter than the age of the stars in almost all cases^{8,30,62}. The inference is that there is a continuous production of small particles in these other systems, similar to the production of interplanetary dust in the Solar System from asteroid collisions.

It is apparent that disks are associated with many stars from birth until they reach main sequence. Evidently, disks evolve in the sense of losing most of their mass in gas and small particles. The planetary formation process sketched in Fig. 2 suggests that residual material will also be cleared from disk inner regions as they age. This clearing has taken place in Vega and β Pic and perhaps in some of the younger systems, as well. Extensive observations of stars intermediate in age between very young (Orion, Taurus) and main sequence (Vega) should fill in this evolutionary hypothesis.

The circumstantial evidence for planet formation is, therefore, quite strong. Most of the observed young disks could create planetary systems like our own. Many would be virtually unobservable by the time their stars were as old as the Sun. The infrared emission from the known interplanetary dust particles around the Sun, the zodiacal light, could not be detected if observed from the distances of the nearest stars³⁰, and the planets themselves have far too little surface area to emit or reflect radiation that could be detected with available instruments. And a significant number of mature stars retain tenuous remnants of their (presumably) more massive disks. In view of the diversity of planets now detected as companions to other stars it is not surprising that the source of planets—disks—are themselves quite diverse.

A clear demonstration that the disks produce planets requires detection of unusually large particles or rocks in the young disks, a measurement of the matter distribution showing inhomogeneities developing as larger bodies are created, and ultimately, the detection of young planets in a few of the older disks. All three approaches are within reach of anticipated instruments in the next 20 years or so.

The strongest way to demonstrate the link between disks and planet production would be detection of planets themselves within the disks. One of the techniques that have been used to detect planets near other stars, the precise tracking of periodic Doppler shifts, is now sensitive enough to detect planets in some of the nearby star-disk systems. However, stellar motion will respond to planets only when one or two planets dominate the gravity field of the disk, and the disks are sufficiently massive in their early, easily detectable phase to diminish the effect of the planets' gravity. Intermediate-aged star-disk systems, in which the disk mass is substantially decreased and most of the mass is in a few planets, are more promising as targets for planet searches, suggesting a future avenue for research into the formation of planetary systems.

Detection of low-mass objects

Not everyone agrees on a precise definition of an extra-solar planet. Stars with masses just below the nuclear burning limit of $0.08 M_{\odot}$ are known as 'brown dwarfs', and most workers agree that brown dwarfs include gaseous objects between about 0.01 and $0.08 M_{\odot}$. The lowest-mass brown dwarfs should be $0.01 M_{\odot}$, the minimum mass of a cloud fragment undergoing gravitational collapse⁶⁴. Jupiter's mass is $0.001 M_{\odot}$, an order of magnitude below the brown-dwarf range. Solid planets such as Earth are

orders of magnitude smaller than any object in the brown-dwarf range. We adopt the view that planets are born from disks, encompassing masses between $\sim 10^{-7}$ (less than Pluto) and $0.01 M_{\odot}$. Objects born in disks should have nearly circular orbits. Under this view, stellar companions near the brown-dwarf boundary of $0.01 M_{\odot}$ would be called planets if they have small orbital eccentricity, and brown dwarfs if the eccentricity is large, such as would be the case if the companions were born as separate cloud fragments during the initial collapse. The companion to 70 Vir, already close to brown-dwarf mass, has an eccentric orbit³ suggesting that it was not born in a disk, but the object near 47 UMa has a circular orbit as expected for a planet; the semi-major axis of the companion to 51 Peg is too small to use eccentricity as a discriminant.

Low-mass objects are extraordinarily difficult to observe. A star produces its light by the nuclear fusion of its large reservoir of hydrogen. Except for a brief period during which the primordial deuterium is burned, objects less massive than $0.08 M_{\odot}$ never become hot enough to trigger hydrogen fusion reactions⁶⁵—they are born warm, with heat derived from the release of gravitational energy as they collapse from extended clouds, and cool slowly generating little radiation. In principle, this cooling period could be observed for objects between about 0.01 and $0.08 M_{\odot}$, but there is little hope of detecting radiation from lighter objects with existing instruments.

Currently, the best approach to finding planets relies on secondary effects: the perturbation of a star's motion or the enhancement or diminution of its light by the planet. The objects recently identified as planets were detected from studies of periodic motion in the velocities of the stars about which they orbit. The apparent position of the star on the sky will exhibit corresponding periodic changes, which can be observed using techniques designed to measure very tiny angular motion (astrometry). It was recently recognized that planets passing near the line of sight to distant stars will cause a brief brightening of the stars owing to gravitational lensing⁶⁶, and this technique holds great promise for future planetary searches. Looking for changes in stellar brightness as a planet passes directly in front of the star, as it will periodically when the orbital inclination is 90° , has also been advocated. However, the subtlety of the effect, and the likely systematic noise from brightness fluctuations intrinsic to the star and complicated by the Earth's atmosphere, have thus far limited the pursuit of this technique⁶⁷. Large planets may even be seen in images taken with large telescopes equipped with improved adaptive optics systems⁶⁸.

Figure 6 compares the sensitivities of the current state-of-the-art techniques with the known low-mass companions in the planetary range. Many hundreds of stars have been searched for radial velocity and angular variations to assess the distribution of masses among companion stars⁶⁹. It is interesting to note that brown dwarfs are not commonly seen as companions to normal stars⁷⁰. They are just beginning to show up in radial velocity surveys: a few per cent of the stars in the large survey by Mayor and his collaborators have companions in the brown-dwarf range (M. Mayor, personal communication, and ref. 71), but it is unlikely that they can account for more than 10% of the companions.

Few, if any, convincing cases for low-mass objects are seen as gravitational 'micro-lensing' events⁷²⁻⁷⁴, even though the surveys for micro-lensing are specifically designed to reveal brown dwarfs. Modest improvements in the technique will make it possible to detect $10^{-6} M_{\odot}$, thus bringing terrestrial planets within the range of observation. Deep images with the Hubble Space Telescope should also have uncovered many brown dwarfs if they constitute a significant fraction of the dark matter in the Galaxy, but none were found⁷⁵.

Although a number of excellent candidates for brown dwarfs are known^{71,76-79}, the preponderance of evidence at present suggests they are not a dominant channel for star formation. Nevertheless, there may well be a large number of planets, even if there are relatively few brown dwarfs. As we have pointed out, planets

and brown dwarfs are created in different processes. The Doppler tracking observations of stars have just reached the range where planets can be detected. In contrast, low-mass objects and long-period orbits are still outside the range of the most sensitive studies.

Because the conditions for planetary-system formation appear to be common among young stars, we believe there are many planetary systems, especially if runaway growth to planets is the normal evolutionary path for these disks. Because the growth timescales of 10^4 – 10^5 years are considerably shorter than the disk lifetimes of more than 10^6 years, we suppose that runaway growth is likely. The mass of heavy elements in the young disks is about the same as that in the planets of the Solar System; there is plenty of material to create planetary systems like our own.

The existence of the newly discovered companions around Sun-like stars suggests that it is even easier to create planets than was first thought. These objects are rather massive for planets, 0.5 – $2M_{\text{Jupiter}}$ for the companion² to 51 Peg and 3.5 – $10M_{\text{Jupiter}}$ for the companion⁴ to 47 UMa (Fig. 6), but probably represent the extremes that are now just detectable in the Doppler tracking surveys. Because massive planets with small orbital radii impose the largest Doppler shifts on the stellar spectra, they are the easiest to find. It is likely that these planets are mostly gaseous. Few of the known disks have enough heavy elements to create such massive planets out of rocks alone, and these relatively massive objects are much closer to their parent stars than Jupiter is to the Sun. They were unanticipated by any theories of planetary-system formation before their discovery⁶⁰.

If we are correct, the future of planetary searches is bright indeed. An interesting issue is whether giant gas planets arise often from these disks—the current limited evidence suggests that

gas is depleted relatively quickly, possibly frustrating assembly of Jupiter-sized bodies. It will be important to build up extensive records to uncover planets several AU from the stars, and it will be very important to push gravitational micro-lensing sensitivities into the terrestrial-planet range: the disks certainly have enough solids to create many terrestrial-sized planets, even if the gas is depleted.

More speculatively, comparison of the properties of planetary systems (particularly the radial distribution of mass) with disks may give insight into the process of assembly, although it could be many years before observations are sensitive enough to produce meaningful results. (We assume here that more than one planet will typically be found around each star.) Both the star-disk systems and the star-planet systems provide opportunities to test our understanding of the origins of the Solar System. Many features of the early stages of planetary growth cannot easily be gleaned from studies of our own planets.

Our inference, that the preponderance of disks suggests planetary systems are common, gives even greater motivation to continue the search for the planetary systems themselves and offers support to what many have thought all along⁸¹, that planetary systems are abundant in the Galaxy. We speculate that if life arises readily on terrestrial planets, then life, too, may be abundant. The recent announcement⁸² that rocks from Mars may contain evidence of life would, if confirmed, support this speculation. Our nearest neighbours may be very near indeed. □

Steven Beckwith is at the Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany; Anneila Sargent is at the Owens Valley Radio Observatory 105-24, California Institute of Technology, Pasadena, California 91125, USA.

1. Wolszczan, A. & Frail, D. A. *Nature* **355**, 145–147 (1992).
2. Mayor, M. & Queloz, D. *Nature* **378**, 355–359 (1995).
3. Marcy, G. W. & Butler, R. P. *Astrophys. J.* **464**, L147–L151 (1996).
4. Butler, R. P. & Marcy, G. W. *Astrophys. J.* **464**, L153–L156 (1996).
5. Safronov, V. in *Evolution of the Protoplanetary Cloud and the Formation of the Earth & Planets* (Rep. No. TTF-667, NASA, Washington DC, 1969). (transl. from Russian).
6. Wetherill, G. W. *Science* **253**, 535–538 (1991).
7. Elsässer, H. & Staude, H. *J. Astron. Astrophys.* **70**, L3–L6 (1978).
8. Aumann, H. H. et al. *Astrophys. J.* **278**, L23–L27 (1983).
9. Smith, B. A. & Terrie, R. *Science* **226**, 1421–1424 (1984).
10. Beckwith, S. et al. *Astrophys. J.* **309**, 755–761 (1986).
11. Adams, F. C., Lada, C. H. & Shu, F. *Astrophys. J.* **308**, 836–853 (1986).
12. Sargent, A. I. & Beckwith, S. *Astrophys. J.* **323**, 294–305 (1987).
13. Cohen, M., Emerson, J. P. & Beichman, C. A. *Astrophys. J.* **339**, 455–473 (1989).
14. Strom, K. M., Strom, S. E., Edwards, S., Cabrit, S. & Skutskie, M. F. *Astron. J.* **97**, 1451–1470 (1989).
15. Beckwith, S. V. W., Sargent, A. I., Chini, R. & Güsten, R. *Astron. J.* **99**, 924–945 (1990).
16. Wood, J. A. & Morfill, G. E. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 329–347 (Univ. Arizona Press, Tucson, 1988).
17. Shu, F. H., Adams, F. C. & Lizano, S. *Annu. Rev. Astron. Astrophys.* **25**, 23–81 (1987).
18. Bertout, C. *Annu. Rev. Astron. Astrophys.* **27**, 351–395 (1989).
19. Bodenheimer, P. *Annu. Rev. Astron. Astrophys.* **33**, 199–238 (1995).
20. Cameron, A. G. W. *Annu. Rev. Astron. Astrophys.* **26**, 441–472 (1988).
21. Shu, F., Najita, J., Galli, D., Ostriker, E. & Lizano, S. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J.) 3–45 (Univ. Arizona Press, Tucson, 1993).
22. Weidenschilling, S. J. *Gerlands Beitr. Geophys.* **96**, 21–33 (1987).
23. Mizuno, H., Markiewicz, W. & Volk, H. *Astron. Astrophys.* **195**, 183–192 (1988).
24. Weidenschilling, S. J. & Cuzzi, J. N. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J.) 1031–1060 (Univ. Arizona Press, Tucson, 1993).
25. Wetherill, G. W. & Stewart, G. R. *Icarus* **106**, 190–209 (1993).
26. Mizuno, H. *Prog. Theor. Phys.* **64**, 544–557 (1980).
27. Stevenson, D. J. *Planet Space Sci.* **30**, 755–764 (1982).
28. Wetherill, G. *Annu. Rev. Earth Planet. Sci.* **18**, 205–256 (1990).
29. Schopf, J. W. in *Major Events in the History of Life* (eds Trần Thanh Vân, J., Trần Thanh Vân, K., Mounolou, J. C., Schneider, J. & McKay, C.) 235–262 (Frontières, Paris, 1992).
30. Backman, D. & Paresce, F. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J.) 1253–1304 (Univ. Arizona Press, Tucson, 1993).
31. Adams, F. C., Lada, C. J. & Shu, F. *Astrophys. J.* **326**, 865–883 (1988).
32. O'Dell, C. R. & Wen, Z. *Astrophys. J.* **436**, 194–202 (1994).
33. McCaughrean, M. J. & O'Dell, C. R. *Astron. J.* **111**, 1977–1986 (1996).
34. Burrows, C. et al. *Astrophys. J.* (in the press).
35. Osterloh, M. & Beckwith, S. V. W. *Astrophys. J.* **439**, 288–302 (1995).
36. Andre, P. & Montmerle, T. *Astrophys. J.* **420**, 837–862 (1994).
37. Dutrey, A., Guilloteau, S. & Simon, M. *Astron. Astrophys.* **286**, 149–159 (1994).
38. Saito, M., Kawabe, R., Ishiguro, M., Miyama, S. M. & Hayashi, M. *Astrophys. J.* **453**, 384–392 (1995).
39. Zuckerman, B., Forveille, T. & Kastner, J. H. *Nature* **373**, 494–496 (1995).
40. Skutskie, M. F. et al. *Astron. J.* **102**, 1749–1752 (1991).
41. Koerner, D. W. & Sargent, A. I. *Astron. J.* **109**, 2138–2145 (1995).
42. Walter, F. M., Brown, A., Mathieu, R. D., Myers, P. C. & Vrba, F. J. *Astron. J.* **96**, 297–325 (1988).
43. Hartmann, L. W., Jones, B. F., Stauffer, J. R. & Kenyon, S. J. *Astron. J.* **101**, 1050–1062 (1991).
44. Lada, C. J., Young, E. T. & Greene, T. P. *Astrophys. J.* **408**, 471–483 (1993).
45. Ladd, E. F., Lada, E. A. & Myers, P. C. *Astrophys. J.* **410**, 168–178 (1993).
46. Strom, K. M., Kepner, J. & Strom, S. E. *Astrophys. J.* **438**, 813–829 (1995).
47. Stauffer, J. R., Prosser, C. F., Hartmann, L. & McCaughrean, M. J. *Astron. J.* **108**, 1375–1381 (1994).
48. Jensen, E. L. N., Mathieu, R. D. & Fuller, G. A. *Astrophys. J.* **429**, L29–L32 (1994).
49. Lay, O. P., Carlstrom, J. E., Hills, R. E. & Phillips, T. G. *Astrophys. J.* **434**, L75–L78 (1994).
50. Mundt, L. G., Looney, L. W., Erickson, W., Grossman, A. & Welch, W. J. *Astrophys. J.* (in the press).
51. Sargent, A. I. & Beckwith, S. V. W. *Astrophys. J.* **382**, L31–L35 (1991).
52. Lagage, P. O. & Pantin, E. *Nature* **369**, 628–630 (1994).
53. Skutskie, M. F., Dutkevitch, D., Strom, S. E., Edwards, S. & Strom, K. M. *Astron. J.* **99**, 1187–1195 (1990).
54. Strom, S. E., Edwards, S. & Skutskie, M. F. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J.) 837–866 (Univ. Arizona Press, Tucson, 1993).
55. Marsh, K. A. & Mahoney, M. J. *Astrophys. J.* **395**, L115–L118 (1992).
56. Boss, A. P. & Yorke, H. W. *Astrophys. J.* **411**, L99–L102 (1993).
57. Beckwith, S. V. W. & Sargent, A. I. *Astrophys. J.* **381**, 250–258 (1991).
58. Mannings, V. & Emerson, J. *Mon. Not. R. Astron. Soc.* **267**, 361–378 (1994).
59. Ossenkopf, V. & Henning, T. *Astron. Astrophys.* **291**, 943–959 (1994).
60. Agladze, N. I., Sievers, J., Jones, S. A., Burlich, J. M. & Beckwith, S. V. W. *Nature* **372**, 243–245 (1995).
61. Pollack, J. B. et al. *Astrophys. J.* **421**, 615–639 (1994).
62. Zuckerman, B. & Becklin, E. E. *Astrophys. J.* **414**, 793–802 (1993).
63. Knacke, R. F. et al. *Astrophys. J.* **418**, 440–450 (1993).
64. Boss, A. P. *Astrophys. J.* **331**, 370–376 (1988).
65. Burrows, A. & Liebert, J. *Rev. Mod. Phys.* **65**, 301–336 (1993).
66. Paczynski, B. *Astrophys. J.* **304**, 1–5 (1986).
67. Parry, I. R., Watson, F. G. & Esperanza Carrasco, B. in *The Future Utilization of Schmidt Telescopes* (eds Chapman, J. et al.) 124–128 (Conf. Ser. 84, Astron. Soc. Pacif., San Francisco, 1995).
68. Angel, J. R. P. *Nature* **368**, 203–207 (1994).
69. Duquenois, A. & Mayor, M. *Astron. Astrophys.* **248**, 485–524 (1991).
70. Marcy, G. & Butler, R. P. in *The Bottom of the Main Sequence—and Beyond* (ed. Tinney, C. G.) 99–108 (Springer, Heidelberg, 1994).
71. Latham, D. W., Mazeh, T., Stefanik, R. P., Mayor, M. & Burki, G. *Nature* **339**, 38–40 (1989).
72. Alcock, C. et al. *Nature* **365**, 619–622 (1993).
73. Aubourg, E. et al. *Nature* **365**, 623–626 (1993).
74. Udalski, A. et al. *Astrophys. J.* **426**, L69–L72 (1994).
75. Bahcall, J. N., Flynn, C., Gould, A. & Kirhakos, S. *Astrophys. J.* **435**, L51–L54 (1994).
76. Stauffer, J. et al. *Astron. J.* **108**, 155–159 (1994).
77. Rebolo, R., Zapatero Osorio, M. R. & Martin, E. L. *Nature* **377**, 129–131 (1995).
78. Nakajima, T. et al. *Nature* **378**, 463–465 (1995).
79. Basri, G., Marcy, G. W. & Graham, J. R. *Astrophys. J.* (in the press).
80. Guillot, T., Burrows, A., Hubbard, W. B., Lunine, J. L. & Saumon, D. *Astrophys. J.* **459**, L35–L38 (1996).
81. Shklovskii, I. S. & Sagan, C. *Intelligent Life in the Universe* (Holden Day, San Francisco, 1996).
82. McKay, D. F. et al. *Science* **273**, 924–930 (1996).

ACKNOWLEDGEMENTS. We profited from discussions with D. Black, D. Koerner, G. Marcy, M. Mayor, M. Meyer, G. Morfill, R. Mundt, D. Padgett, K. Stapelfeldt, F. Shu and H. Volk. We thank the referees for their many suggestions that improved the original manuscript.

CORRESPONDENCE should be addressed to S.V.W.B. (e-mail swb@mpia-hd.mpg.de).

Macroscopic quantum tunnelling of magnetization in a single crystal of nanomagnets

L. Thomas*, F. Lioni*, R. Ballou*, D. Gatteschi†, R. Sessoli† & B. Barbara*

* Laboratoire de Magnétisme Louis Néel, CNRS, BP 166, 38042 Grenoble, Cedex 9, France

† Department of Chemistry University of Firenze, 50144, Italy

The precise manner in which quantum-mechanical behaviour at the microscopic level underlies classical behaviour at the macroscopic level remains unclear, despite seventy years of theoretical investigation. Experimentally, the crossover between these regimes can be explored by looking for signatures of quantum-mechanical behaviour—such as tunnelling—in macroscopic systems¹. Magnetic systems (such as small grains, spin glasses and thin films) are often investigated in this way^{2–12} because transitions between different magnetic states can be closely monitored. But transitions between states can be induced by thermal fluctuations, as well as by tunnelling, and definitive identification of macroscopic tunnelling events in these complex systems is therefore difficult¹³. Here we report the results of low-temperature experiments on a single crystal composed of superparamagnetic manganese clusters ($\text{Mn}_{12}\text{-ac}$), which clearly demonstrate the existence of quantum-mechanical tunnelling of the bulk magnetization. In an applied magnetic field, the magnetization shows hysteresis loops with a distinct 'staircase' structure: the steps occur at values of the applied field where the energies of different collective spin states of the manganese clusters coincide. At these special values of the field, relaxation from one spin state to another is enhanced above the thermally activated rate by the action of resonant quantum-mechanical tunnelling. These observations corroborate the results of similar experiments performed recently on a system of oriented crystallites made from a powdered sample⁴.

A single crystal of $\text{Mn}_{12}\text{-ac}$ can be considered as a single array of nanomagnets with the same directions of magnetic anisotropy with respect to the applied field. As such, it has the advantage of having a single level scheme corresponding to that of one nanomagnet of macrospin S . A further advantage of using a single crystal is that the applied field becomes a control parameter than can potentially be used to tune the quantum tunnelling of magnetization. Further, dipolar interactions between nanomagnets are well defined and can be calculated accurately for a single grain. The crystals of $\text{Mn}_{12}\text{-ac}$ are built of discrete dodecanuclear $[\text{Mn}_{12}(\text{CH}_3\text{COO})_{16}(\text{H}_2\text{O})_4\text{O}_{12}] \cdot 2\text{CH}_3\text{COOH} \cdot 4\text{H}_2\text{O}$ molecules, water of crystallization and disordered acetic acid molecules¹⁴. The dodecanuclear molecule has an overall tetragonal symmetry and the unit-cell volume is $3716 \text{ \AA}^3$. The 12 manganese ions of each unit cell form a mixed valence cluster with 4 inner Mn^{4+} ions (of spin $S = 3/2$) surrounded by 8 Mn^{3+} ions (of spin $S = 2$). Exchange interactions inside the cluster stabilize a ferrimagnetic cluster ground state of collective spin $S = 10$ (parallel alignment of the Mn^{3+} and Mn^{4+} individual spins, and antiparallel alignment of the Mn^{3+} and Mn^{4+} spins)¹⁵. A strong anisotropy energy of uniaxial symmetry, DS_z^2 with $D = 0.61 \text{ K}$ per cluster, is induced by the crystalline electric field¹⁶. Dipolar interactions between the clusters have been estimated to be about 0.01 T from numerical micromagnetic calculations, assuming dipoles with spins $S = 10$ distributed in the crystal structure of $\text{Mn}_{12}\text{-ac}$.

Magnetization measurements are performed on a single crystal of parallelepiped shape with volume $V = 0.066 \pm 0.008 \text{ mm}^3$, at 12 different temperatures between 1.55 K and 3.2 K , in applied

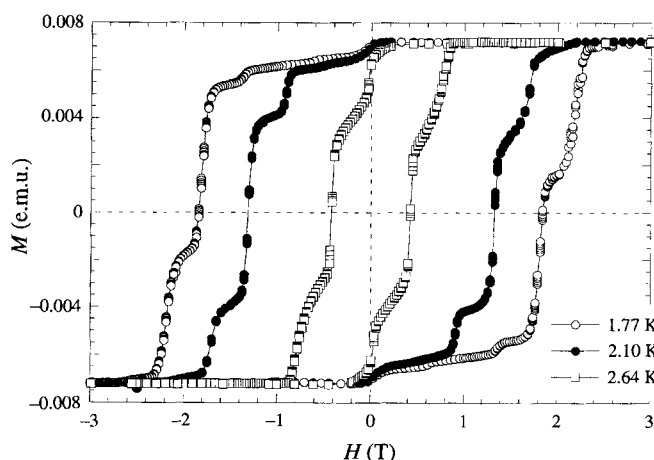


FIG. 1 Magnetization hysteresis loops measured on the 0.066 mm^3 single crystal at 1.77 K , 2.10 K and 2.64 K between magnetic fields of 5 T and -5 T , with a SQUID magnetometer. The sensitivity allowed the detection of magnetization variations of $\sim 0.5\%$ of the saturation magnetization M_s , of our crystal. The magnetic field was applied along the magnetization axis, which coincides with the longest dimension of the crystal. The quality of its orientation with respect to the applied field was checked by comparing the remanent magnetization M_r at low temperature, with M_s (in our case $M_r/M_s > 99\%$). Before starting the experiment, the sample was cooled from above the blocking temperature in an applied of 5 T . The hysteresis loop is determined by the following method, which fixes the characteristic time windows of the measurements: when a measurement at a given field has ended, the field is changed during 10 s to reach the next value; after a pause of 400 s , to allow stabilization, the measurements at the new value of the field are taken (by an axial extraction method between two measuring coils). As more than one extraction is performed (each one is 60 s in duration), the waiting time at each field value is about 10 min . Within the time window used for the measurements, the fast magnetization variations of each step are observed with a field increment fixed at about 0.025 T .

magnetic field strengths up to 5 T . A 'staircase' hysteresis loop is discovered, with a succession of steep and flat regions, when the field direction is opposite to the magnetization (Fig. 1). In the flat regions of the hysteresis loop, the magnetization relaxation times are much larger than the measurement time window ($\sim 600 \text{ s}$). In the steep regions, the relaxation times are of the order of the measurement time window. The plot of $\partial M / \partial H$ against H allows us to define accurately the critical fields H_n at which the relaxation times are at a minimum, as well as the width ΔH_n of the relaxations from long to short relaxation times (see for example, inset of Fig. 2). The values of H_n and ΔH_n found in a range of temperature and magnetic field strength ($1.55 \text{ K} < T < 3.2 \text{ K}$ and $-5 \text{ T} < H < 5 \text{ T}$) are summarized in Fig. 2. There are two main observations: first, the jumps in magnetization are regularly spaced in H , with an anomaly at the lowest temperatures where a few jumps are not observed; second, the $\partial M / \partial H$ versus H transitions have a lorentzian shape with widths ΔH_n well above the experimental resolution. Measurements of magnetic relaxation were also performed in different conditions (Fig. 3). The relaxation time τ oscillates with respect to the applied field. It decays monotonically as the applied field is increased with deep minima at the field strengths where there is a step in the magnetic hysteresis loop. It is interesting to note that the two sets of relaxation times in Fig. 3, corresponding to the steep and flat portions of the hysteresis loop, are very close to those previously observed on orientated grains of the same material.³

Our data shows that regardless of the measured widths ΔH_n and small bias associated with each H_n , the same sequence of magnetization jumps, from $H_0 = 0$ to $H_6 = 2.64 \text{ T}$, with a constant

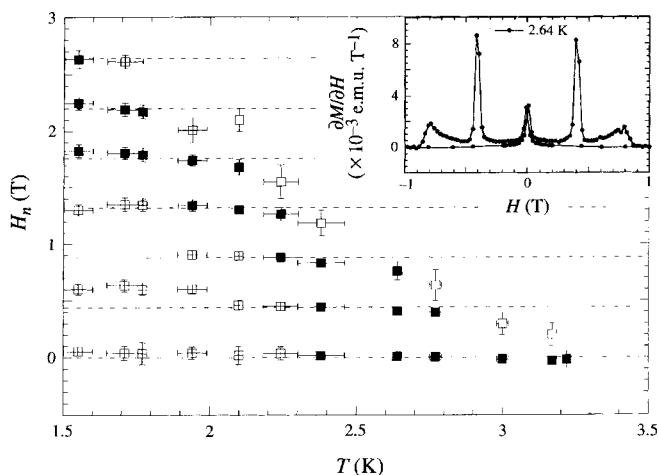


FIG. 2 Fields H_n for which the jumps are observed are plotted versus temperature. At the lowest temperatures (1.55 K and 1.70 K), we observe six jumps. The number of jumps decreases as the temperature increases (for example four jumps at 2.4 K, three jumps at 2.6 K and only one jump at 3.2 K in zero field when the relaxation is too fast and T is close to the blocking temperature). The first magnetization jump at $H_0 = 0$ (and the first minimum of $\tau(H)$), arises from the coincidence of the ten states of the symmetrical, double-well level scheme (+10 and -10, +9 and -9, ..., +1 and -1); the magnetization jump at $H_1 = 0.44$ T (and the second minimum of $\tau(H)$) corresponds to the coincidence of the 9 levels (+10 and -9, +9 and -8, ..., +2 and -1) and so on. The last jump (and the corresponding $\tau(H)$ minimum) observed at $H_6 = 2.64$ T, arises from four coincidences only (+10 and -4, +9 and -3, +8 and -2, +7 and -1). The horizontal bars account for the temperature stability over the whole duration of the experiments. The vertical bars account for the width ΔH_n associated with a jump (ΔH_n is deduced from the derivative $\partial M/\partial H$ of the hysteresis loop). The filled squares show jumps of more than 10% of the total magnetization. The crossed squares show jumps of less than 10% of the total magnetization. Decays of the magnetization over a range of fields, larger by one order of magnitude than the dipolar field, are shown by the open squares. The inset shows an example of a field variation of $\partial M/\partial H$ (at $T = 2.64$ K). On the main plot, the horizontal dashed lines indicate where level matching occurs.

field increment $H_{n+1} - H_n = \Delta H = 0.44$ T, are observed at each temperature, and only in negative fields (that is when the field is antiparallel to the initial remanent magnetization). This behaviour is the consequence of a succession of level crossings of the collective spin $S = 10$. A crossing of the states $M_S^m = -m$ and $M_S^{m+n} = -m + n$ occurs when the field H_n is such that $-Dm^2 + g\mu_B m H_n = -D(m-n)^2 - g\mu_B (m-n)H_n$, which gives $g\mu_B H_n = nD$, where D is the uniaxial anisotropy energy. Since the fields H_n are independent of m , the level coincidence occurs simultaneously for all the states, except those with lowest energies ($m \geq S - n$). Quantitatively, our transition-field increment, $H_{n+1} - H_n = \Delta H = D/g\mu_B = 0.44$ T, gives $D = 0.60$ K, which is close to the accepted value of 0.61 K per cluster¹⁶.

It is therefore clear that the magnetic relaxation of the Mn_{12} -ac crystal is strongly enhanced when the spin-up and spin-down level schemes coincide. This is evidence for resonant macroscopic quantum tunnelling in a magnetic system. Similar conclusions have recently been drawn from magnetization measurements performed on powdered Mn_{12} -ac material, mixed into Stycast 1226 and oriented in a magnetic field of 5.5 T at 300 K (ref. 4). The observation of tunnelling effects in magnetic grains, with or without domain walls has also been reported (ref. 2 and references therein). Even when features compatible with quantum tunnelling such as magnetic relaxation plateaux) were measured in several systems, this phenomenon was not proved¹³ because the experi-

ments were not carried out at temperatures low enough to rule out non-quantum effects such as a combination of energy barrier and switching field distributions or self-critical magnetic avalanches^{2,3,5}. Only in few cases have experiments been performed at temperatures below 1 K (on oriented grains of Mn_{12} (refs 3, 6), on Co (ref. 7) and Ni (ref 7, 8) and on ferritin⁹⁻¹²). In all cases, qualitative but not quantitative agreement with the existing theories of quantum tunnelling, was achieved. This is not the case in the present work, where relaxation time minima, observed for particular values of the applied field are directly, and very simply related to the anisotropy energy. This striking evidence for resonant transitions is quite new—the possibility of such transitions had not been raised before the studies on Mn_{12} -ac by Barbara *et al.*³, and more recently by Friedman *et al.*⁴. A more complete description of these observations and, in particular, of the fast variation of the relaxation time with temperature (Fig. 3, inset), requires taking into account the thermal activation within a given quantum well. The decrease of the relaxation time with temperature must be associated with tunnelling from excited levels with smaller Boltzmann populations but larger tunnelling probabilities (owing to smaller barrier height and width). Simple estimates suggest that each relaxation time minimum corresponds to resonant tunnelling from a given level, with a relaxation time of the order of magnitude of the experimental time window. As shown in Fig. 3, these minima (due to resonant tunnelling) are superimposed on a continuous curve which is strongly dependent on temperature (owing to thermal activation). These features are fully consistent with the observation^{3,6} of thermal activation observed between 1 and 2 K and the extremely long relaxation times observed below 1 K.

The theory of tunnelling in the presence of phonons^{17,18} with a strict tetragonal symmetry (as assumed for Mn_{12} -ac) does not

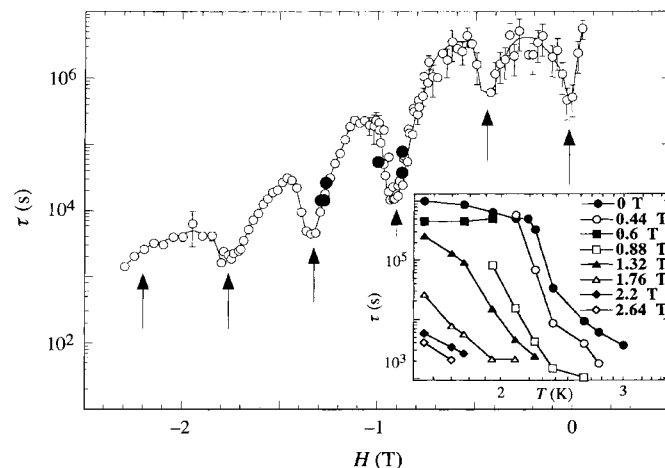


FIG. 3 Relaxation times at temperature 2.10 K versus H , obtained from repeated measurements for given H and T on the hysteresis loop. Use was made of the repeated measurements at each point of a hysteresis loop at short times ($300 \text{ s} < t < 600 \text{ s}$). The data fit very well with a linearized exponential approximation: $M(t) = M_0(1 - t/\tau) - M_s$ from which τ (open circles) is extracted. A careful check of these relaxation times has been made by performing relaxation measurements over ~ 20 – 50 h at various fields and temperatures ($H = -0.33$ T, -1.43 T, -1.67 T and -1.77 T at 1.77 K; $H = -0.87$ T, -1.00 T, -1.26 T at 2.10 K; $H = 0$ T at 2.65 K). Nearly-exponential decays in M were observed, and at the largest times, they tended to the saturation magnetization $-M_s$ parallel to the field direction. This method leads to values of τ (filled circles) in good agreement with those obtained at short timescales. The arrows show the fields at which jumps are expected from the level coincidence scheme. The inset shows the relaxation time drops plotted against temperature in log-log scale, for several applied fields.

predict resonant tunnelling. Relaxation time maxima are found when spin-up or spin-down level schemes coincide, whereas our experiments show the contrary. Moreover, this model involves only transitions with $\Delta m = \pm 4$, in contrast with experiments where almost all of the $\Delta m = \pm 1$ transitions are observed. However, the theory including intermolecular dipolar coupling, where the relevant matrix elements are provided by the environment, predicts that tunnelling transitions can occur, and involve clusters of molecules with minima in the relaxation times at fields for which level matching occurs¹⁹. In this model, the width ΔH_n of these transitions should account for a spatially inhomogeneous broadening of the relaxation time due to dipolar couplings (and perhaps also the coupling of the macrospin system to incoherent phonons and nuclear spins). This agrees with the measured values of ΔH_n which are close to dipolar and hyperfine interactions. In this model, transitions involving more than two molecules are much less probable than pairwise transitions. The fact that these transitions have $\Delta m = \pm 2$, corresponds with the absence of a jump at $H_2 = 8.8$ kOe at temperatures of 1.55 and 1.77 K. The existence of a weak, uniaxial, fourth-order anisotropy term in the spin hamiltonian, might also lead to small deviations to pure $\Delta m = \pm 1$ transitions. Besides single particle thermal activation, environmental noise (such as thermal fluctuations of cluster dipoles) and a decrease of local magnetic fields during the relaxation experiment (resulting from the relaxation of the magnetization on the hysteresis loop) must contribute to temperature-dependence, assisted tunnelling¹.

These experiments give clear evidence for the quantization of the magnetization in the hysteresis loop of the macrospin of a single crystal. This phenomenon is quite different from the reversible steps resulting from level crossings²⁰ in systems with low-spin ground states. In our case quantization of the magnetization comes from the tunnelling (and not spin rotation) of an assembly of coupled nanoscale magnets with integer macrospins²¹, where matrix elements breaking the precessional spin-symmetry result from couplings of each cluster spin to its environment (magnetic noise, incoherent phonons and dipolar and hyperfine interactions). Macroscopic quantum tunnelling for this magnetic system was observed with finite field resolution, owing to the strong, inhomogeneous level broadening which arises from couplings to the environment. This is the reason why it is possible to tune the quantum relaxation at the macroscopic scale, using a quite conventional SQUID magnetometer. $\square$

Received 1 March; accepted 26 July 1996.

- Leggett, A. J. in *Quantum Tunneling of the Magnetisation* (eds Gunther, L. & Barbara, B.) 1–18 (NATO ASI Series E: Applied Sciences-Vol. 301) (Kluwer, Dordrecht, 1995).
- Barbara, B. et al. *J. Appl. Phys.* **73**, 6703–6708 (1993).
- Barbara, B. et al. *J. Mag. Magn. Mater.* **140–144**, 1825–1828 (1995).
- Friedman, J., Sarachik, M. P., Tejada, J., Maciejewski, J. & Ziolo, R. *Phys. Rev. Lett.* **76**, 3820–3833 (1996).
- Barbara, B. et al. in *Studies of Magnetic Properties of Fine Particles and their Relevance to Material Science* (eds Dormann, J. L. & Fiorani, D.) 235–242 (Elsevier, Amsterdam, 1991).
- Paulsen, C. & Park, J.-G. in *Quantum Tunneling of the Magnetisation* (eds Gunther, L. & Barbara, B.) 189–207 (NATO ASI Series E: Applied Sciences-Vol. 301) (Kluwer, Dordrecht, 1995).
- Wernsdorfer, W. et al. *J. Mag. Mag. Mat.* **145**, 33–37 (1995).
- Giordano, N. & Hong, K. in *Quantum Tunneling of the Magnetisation* (eds Gunther, L. & Barbara, B.) 257–272 (NATO ASI Series E: Applied Sciences-Vol. 301) (Kluwer, Dordrecht, 1995).
- Awschalom, D. D., Smyth, J. F., Grinstein, G., Di Vincenzo, D. P. & Loss, D. *Phys. Rev. Lett.* **68**, 3092–3095 (1992).
- Gider, S., Awschalom, D. D., Douglas, T., Mann, S. & Chaparala, M. *Science* **268**, 77–80 (1995).
- Tejada, J. *Science* **272**, 424–424 (1996).
- Garg, A. *Science* **272**, 424–425 (1996).
- Stamp, P. C. E. *Nature* **359**, 366–367 (1992).
- Sessoli, R. et al. *J. Am. Chem. Soc.* **113**, 5873–5874 (1991).
- Caneschi, A. et al. *J. Am. Chem. Soc.* **113**, 5873–5874 (1991).
- Sessoli, R., Gatteschi, D., Caneschi, A. & Novak, M. A. *Nature* **365**, 141–143 (1993).
- Politi, P., Rettori, A., Hartmann-Boutron, F. & Villain, J. *Phys. Rev. Lett.* **75**, 537–540 (1995).
- Prokofev, N. V. & Stamp, P. C. E. *J. Low Temp. Phys.* (in the press).
- Burn, A., Prokofev, N. V. & Stamp, P. C. E. *Phys. Rev. Lett.* **76**, 3040 (1995).
- Taft, K. L. et al. *J. Am. Chem. Soc.* **116**, 823 (1994).
- Chudnovsky, E. M. & Di Vincenzo, D. P. *Phys. Rev. B* **48**, 10548–10551 (1993).

ACKNOWLEDGEMENTS. We thank Andrea Caneschi for sample preparation as well as Stewart Barnes, Nikolas Garcia, Carley Paulsen and Anatolij Zvezdin for discussions. B.B. thanks Tony Leggett for discussions on macroscopic tunnelling in magnets.

CORRESPONDENCE should be addressed to B.B. (e-mail: barbara@labs.polycns-gre.fr).

Synthesis of hydroazafullerene $C_{59}HN$, the parent hydroheterofullerene

Majid Keshavarz-K.*, Rosario González*, Robin G. Hicks*, Gordana Srdanov*, Vojislav I. Srdanov*, Tasha G. Collins*, Jan Cornelis Hummelen*, Cheryl Bellavia-Lund*, James Pavlovich*, Fred Wudl* & Karoly Holczer†

* Departments of Chemistry and Materials, Institute for Polymers and Organic Solids, University of California, Santa Barbara, California 93106-5090, USA

† Department of Physics, University of California, Los Angeles, California 90024, USA

THE electronic and geometric properties of C_{60} can be perturbed by replacing one or more carbon atoms of the fullerene skeleton with an atom of a different element. Exchange of one carbon atom with nitrogen, a trivalent atom with a lone pair of electrons, produces the azafullerene radical $C_{59}N$, which is isoelectronic with the C_{60} radical anion. The process is slightly similar to doping silicon with phosphorus¹. We have previously described the synthesis of the azafullerene dimer²; here we report the bulk preparation of the simplest azafullerene, $C_{59}HN$. The electronic, vibrational and ^{13}C NMR spectroscopic features of $C_{59}HN$ are similar to those of the dimer², except for the signature of the sp^3 (C–H) carbon. $C_{59}HN$ should open the door to a new chemistry of heterofullerenes.

The parent hydroazafullerene, $C_{59}HN$, is of special interest because it is the simplest closed-shell azafullerene and its spectroscopic properties can be useful as a reference for the characterization of more complex derivatives, especially $(C_{59}N)_2$. In addition, it is in principle a precursor for the preparation of other azafullerene derivatives, $C_{59}RN$, (where R is Ph_2CH-), by deprotonation and subsequent quenching of the anion with various electrophiles^{3,4}. It is also a precursor, by deprotonation, to doped $MC_{59}N$ salts (M is Na^+ , K^+ , Rb^+ and so on), which are isoelectronic with M_2C_{60} solids.

The azafullerene radical was synthesized from C_{60} in three steps (Fig. 1) and isolated as its dimer² $(C_{59}N)_2$ (compound 7). The dimer was doped with potassium metal, to $K_6C_{59}N$ which was found to be isostructural (but not isoelectronic) with its buckminsterfullerene analogue K_6C_{60} , with virtually identical lattice dimensions⁵. We have corroborated the previously proposed mechanism² for the formation of $(C_{59}N)_2$, from MEM-substituted ketolactam (1) and strong acid (Fig. 1), by intercepting the

TABLE 1 Redox potentials for heterofullerene species and C_{60}

	$(C_{59}N)_2$	$C_{59}HN$	C_{60}
E_1^{red}	–992	–1,106*	–1,123
	–1,071		
E_2^{red}	–1,424	–1,500*	–1,450
	–1,485		
E_3^{red}	–1,979		–1,913
	–2,089		
E^{ox}	+886*	+823*	+1,300*† +1,260‡

Potentials are in mV relative to the ferrocene/ferrocenium electrochemical couple, in ODCB solvent, unless otherwise stated. E_1^{red} is the reduction potential of the first electrochemical process ($C_{59}HN \rightarrow C_{59}HN^{\cdot-}$); E_2^{red} , $C_{59}HN^{\cdot-} \rightarrow C_{59}HN^{2-}$; E_3^{red} , $C_{59}HN^{2-} \rightarrow C_{59}HN^{3-}$; E^{ox} , $C_{59}HN \rightarrow C_{59}HN^{\cdot+}$.

* Chemically irreversible.

† See ref. 10.

‡ Chemically reversible, in s-tetrachlorethane (see ref. 11).

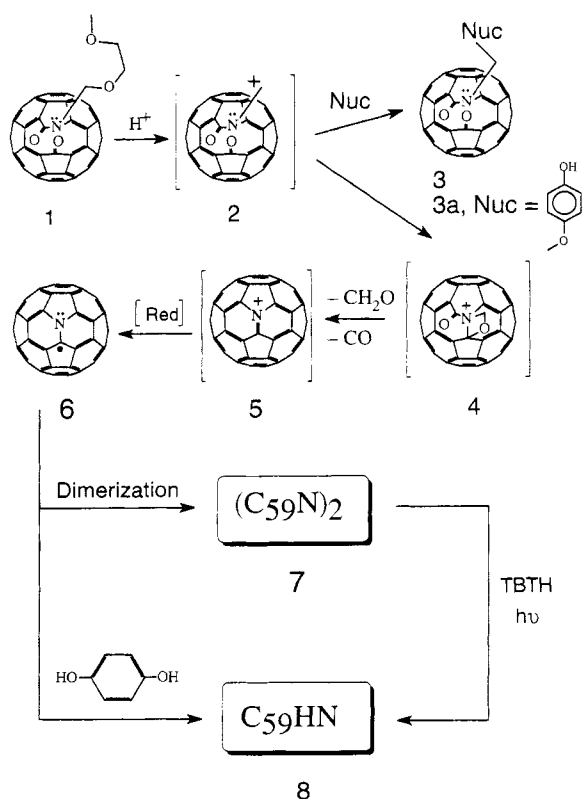


FIG. 1 Schematic representation of proposed chemical transformations involved in the synthesis of $C_{59}HN$ (compound **8**). Nuc, nucleophile; TBTH, tributyltin hydride; [Red], reductive process. When a large excess of hydroquinone is used instead of reduction, trapping of compound **2** affords the ether (**3a**).

dimerization path ($5 \rightarrow 6 \rightarrow 7$) (MEM is methoxyethoxymethyl). A reducing agent stable to strong acid in refluxing 1, 2-dichlorobenzene (ODCB), was required to reduce compound **5** (or **6**), prevent dimerization and give $C_{59}HN$. The ketolactam (**1**) in a solution of excess *p*-toluenesulphonic acid monohydrate and hydroquinone in ODCB at 170–180 °C, in argon, was converted to the desired hydroazafullerene with a high yield, after chromatographic purification. A large excess of hydroquinone (75-fold) traps the intermediate compound **2** and gives the hydroquinone ether (**3a**) as the major product. The product, $C_{59}HN$, was also obtained when $(C_{59}N)_2$ was irradiated through a Kapton filter, in ODCB in the presence of tributyltin hydride (TBTH) in argon gas at 25–35 °C. These reactions can be monitored by HPLC (using a 4.6×10 mm Cosmosil column with a toluene eluant at a flow rate of 1 ml min^{-1} ; ultraviolet detection at 326 nm; retention time is 7.4 min for $C_{59}HN$ and 14.1 min for $(C_{59}N)_2$). A negative-ion electrospray mass spectrum of a toluene solution of the above product shows a base peak at a mass-to-charge ratio $m/z = 723$, confirming that the molecular composition of the isolated product is $C_{59}HN$ (Fig. 2). Desorption chemical ionization (DCI) mass spectra show peaks at 723 (for M) and 722 (for M–H) accompanied by 'shrink-wrap' peaks at m/z 696, 672, 648 and 624. The efficient loss of 26 AMU (loss of CN at $m/z = 696$) and successive 24 AMU fragment loss (loss of C_2 at $m/z = 672, 648$ and so on) is similar to the behaviour of $(C_{59}N)_2$ (ref. 2).

The product, $C_{59}HN$ is green in common solvents (ODCB, carbon disulphide, toluene) and its ultraviolet–visible spectrum has absorptions at wavelengths of 324, 446, 596, 738 and 818 nm, virtually identical to those of $(C_{59}N)_2$ (Fig. 3a). The pure powder diffuse reflectance infrared technique (DRIFT) spectrum (Fig. 3b) shows multiple absorptions in the four regions where C_{60} shows single peaks (at wavenumbers of 527, 576, 1,182 and $1,428 \text{ cm}^{-1}$). In addition, there are three peaks at wavenumbers of 818, 839 and

845 cm^{-1} and multiple weak absorptions between 1,490 and $1,590 \text{ cm}^{-1}$; the strongest peak in this region is at $1,550 \text{ cm}^{-1}$.

The laser Raman scattering, structural examination is much more informative. A comparative study is shown in Fig. 3c. The salient features of this figure are the red-shift of the A_g -derived modes at 1,460 and 491 cm^{-1} and the doubling/broadening of the H_g -derived modes at 1,566 and 269 cm^{-1} owing to symmetry reduction on going from fullerene to heterofullerene.

The ^1H NMR of compound **8** in carbon disulphide solvent (external D_2O lock) has a single resonance of 8.5 p.p.m. The ^{13}C NMR spectrum of compound **8** is very similar to that of compound **7** in the sp^2 fullerene region² (124–157 p.p.m.) but unlike compound **7**, compound **8** shows an intense signal of 72.1 p.p.m. The C_s symmetry, 'closed' structure, compound **8a** (see Fig. 4) can be unambiguously assigned to hydroazafullerene, $C_{59}HN$, on the basis of the following NMR results: the number of carbon resonances in the sp^2 region (30 are expected); the signal in the sp^3 region (72.1 p.p.m.); the proton–carbon coupled NMR experiment ($J_{C-H} = 162.0 \text{ Hz}$); and the comparison of these data with those of 1,2-dihydrofullerenes (see Fig. 5). Because J_{C-H} coupling constants are directly proportional to the degree of *s*-character of a C–H bond, they can be used to assign hybridization to carbon atoms in organic molecules. If the saturated carbon in $C_{59}HN$ were sp^3 -hybridized, a value of 140–145 Hz would be expected for J_{C-H} , but the observed large coupling constant of 162 Hz indicates a higher-than- sp^3 (intermediate between sp^3 and sp^2) *s*-character⁶ compared to a similar bond in $C_{60}HCN$ ($J_{C-H} = 143.4 \text{ Hz}$). Attachment of this carbon to nitrogen is expected to increase J by 5–10 Hz (giving a value of 145–150 Hz) but not the observed increase of 17–22 Hz (J_{C-H} of 162 Hz). Additional structural information could be obtained from sophisticated two-dimensional and three-dimensional pulsed electron spin resonance spectroscopy (FTESR). A pure sample of $C_{59}HN$ (in solution and in the solid state) contains a paramagnetic component which has features which can be attributed to $C_{59}N^{\cdot-}$.

Much like $C_{60}HCN$ (ref. 3) and $C_{60}H_2$ (refs 7, 8), $C_{59}HN$ exhibits only quasi-reversible reduction waves and an irreversible oxidation wave (see Fig. 6 and Table 1). The second wave appears to be more reversible than the first. An explanation for this observation depends crucially on the loss of a hydrogen atom from the radical anion of these systems as shown (for $C_{59}HN$) in Fig. 6. The remarkable similarity of the spectroscopic properties of compounds **7** and **8** allowed us to describe the structure of compound **7** more accurately than before². In the previous attempt², the absence of an NMR signal due to the sp^3 -hybridized carbons did not allow a distinction between compounds **7a** and **7b**. Because the NMR evidence presented above demands a closed structure

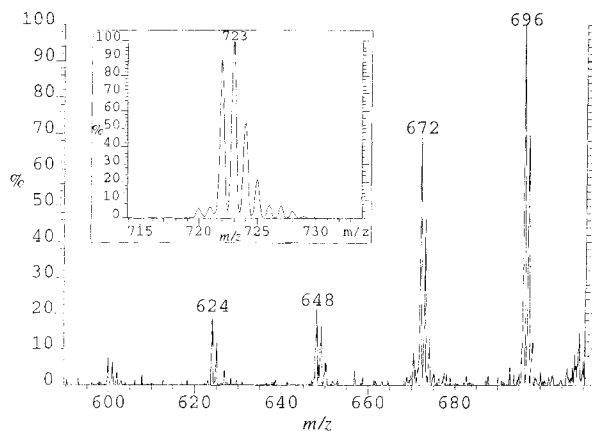


FIG. 2 Section of DCI mass spectrum showing the 'shrink-wrap' pattern. The inset is an expansion around the peak at mass-to-charge ratio $m/z = 723$. The expanded pattern gives a good fit to the calculated isotope mass ratios.

FIG. 3 *a*, Ultraviolet-visible spectrum of $C_{59}HN$ (dotted line) and $(C_{59}N)_2$ (solid line) under the same conditions in toluene. *b*, DRIFT (diffuse reflectance infrared technique) spectrum of pure $C_{59}HN$. *c*, Fourier transform-Raman spectra of $C_{59}HN$ (top) and of C_{60} (bottom) obtained with 1,064-nm laser line. The background noise (top) is due to broad luminescence (see *d*) in the mid-infrared region of $C_{59}HN$ (absent in C_{60}), which prevents observation of weak Raman modes that are expected to appear in the symmetry-reduced molecule. The sharpest features at wavenumbers of 1,460 and 491 cm^{-1} in the $C_{59}HN$ spectrum are red-shifted by 7 and 3 cm^{-1} , respectively, with respect to the two A_g modes of C_{60} , while the $C_{60}H_g$ parentage of the broader features centred at wavenumbers 1,566, 1,091, 760 and 269 cm^{-1} is evident. The red-shift with respect to the parent C_{60} modes is in part due to the greater mass of $C_{59}HN$. *d*, Luminescence of $C_{59}HN$ in the region 10,000–6,000 cm^{-1} (absolute wavenumbers). The 1,064-nm excitation Nd-YAG line appears as a sharp line at 9,398 cm^{-1} .

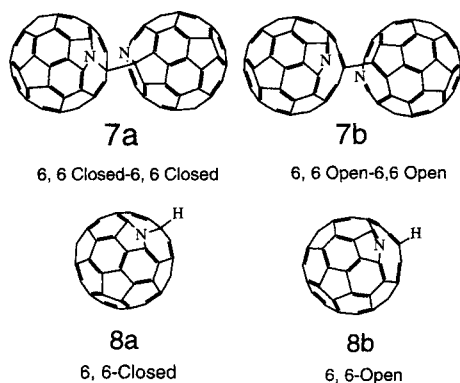
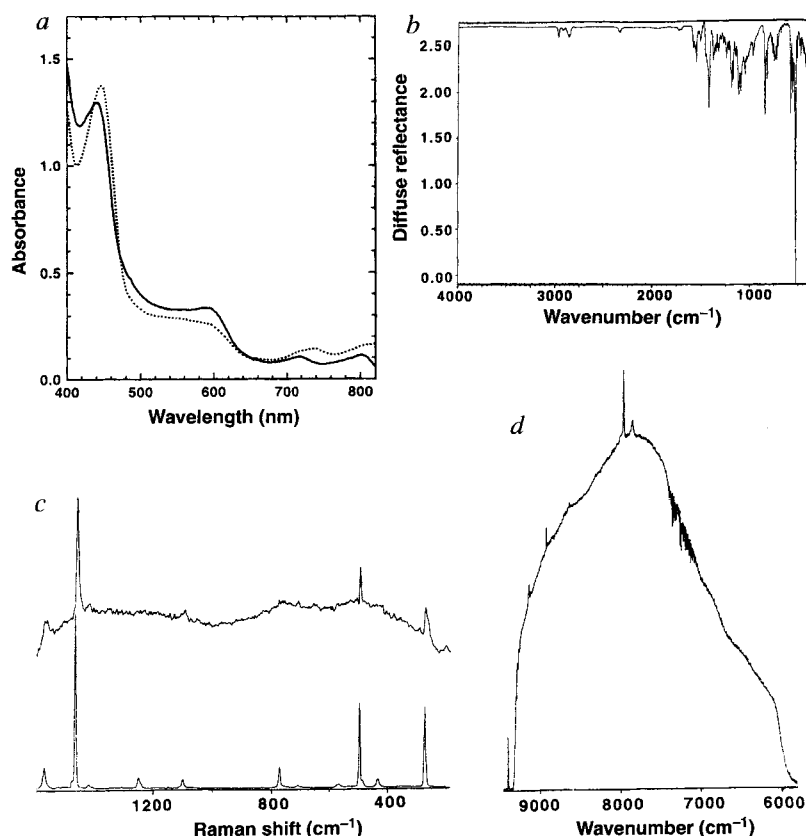


FIG. 4 Open and closed structures of $(C_{59}N)_2$ and $C_{59}HN$. It is shown by spectroscopy that the closed forms (**7a** and **8a**) are the correct structures.

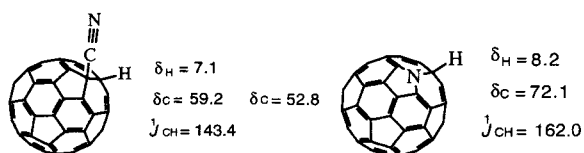


FIG. 5 Selected NMR data of dihydrofullerene $C_{60}HCN$ (CS_2 solvent with D_2O external lock) is shown for comparison with $C_{59}HN$ (see ref. 3).

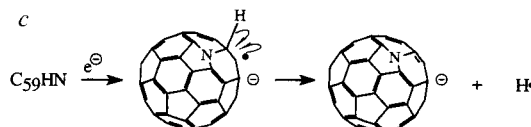
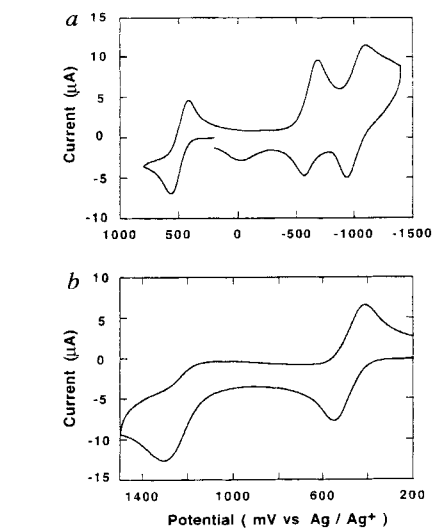


FIG. 6 *a*, *b*, Cyclic voltammograms of $C_{59}HN$. *a*, Reduction processes; *b*, oxidation processes. The redox process at $\sim +500$ mV in both panels is the ferrocene/ferrocenium redox couple. Electrolyte, TBA BF_4 /ODCB; Pt working and counter electrodes; scan rate, 100 $mV s^{-1}$. *c*, Schematic representation of hydrogen-atom loss upon addition of one electron to $C_{59}HN$.

for compound **8**, and the electronic absorption spectrum of compound **7** is qualitatively the same as that of **8**, the closed-closed structure (**7a**) can be assigned to compound **7** as well.

Unlike C_{60} , hydroazafullerene has an acidic C–H bond. Exposure to base such as potassium *t*-butoxide should produce the azafulleride anion, isoelectronic with C_{60}^{2-} . This property will make the functionalization chemistry of heterofullerenes different from that of the parent. Another aspect of $C_{59}N^-$ is its ground state, which could be a triplet, in analogy with C_{60}^{2-} (ref. 9). The $A^+C_{59}N^-$ solids can be further doped to $A_2C_{59}N$ which are isoelectronic with the A_3C_{60} -superconductors. Unlike the latter, which are pyrophoric, the former should be much more stable to exposure to the atmosphere—an important step towards technological applications. □

Received 10 June; accepted 30 July 1996.

- Guo, T., Jin, C. & Smalley, R. E. *J. Phys. Chem.* **95**, 4948–4950 (1991).
- Hummelen, J. C., Knight, B., Pavlovich, J., González, R. & Wudl, F. *Science* **269**, 1554–1556 (1995).
- Keshavarz-K, M., Knight, B., Srdanov, G. & Wudl, F. *J. Am. Chem. Soc.* **117**, 11371–11372 (1995).
- Kitagawa, T., Tanaka, T., Takata, Y., Takeuchi, K. & Komatsu, K. *J. Org. Chem.* **60**, 1490–1491 (1995).
- Prassides, K. et al. *Science* **271**, 1833–1835 (1996).
- Andreoni, W., Curioni, A., Keshavarz-K, M. & Wudl, F. *J. Am. Chem. Soc.* (in the press).
- Niyazymbetov, M. E., Evans, D. H., Lerke, S. A., Cahill, P. A. & Henderson, C. C. *J. Phys. Chem.* **98**, 13093–13098 (1994).
- Guarr, T. F., Meier, M. S., Vance, V. K. & Clayton, M. J. *Am. Chem. Soc.* **115**, 9862–9863 (1993).
- Boyd, P. D. W. et al. *J. Am. Chem. Soc.* **117**, 2907–2914 (1995).
- Dubois, D. & Kadish, K. J. *Am. Chem. Soc.* **113**, 7773–7774 (1991).
- Xie, Q., Anois, F. & Echegoyen, L. J. *Am. Chem. Soc.* **115**, 9818–9819 (1993).

ACKNOWLEDGEMENTS. We thank the National Science Foundation for support through the Materials Research Laboratory as well as individual investigator grants. R. G. thanks the Ministerio de Educación y Ciencia (Spain) for a postdoctoral fellowship. R.G.H. thanks the National Research Council of Canada for a postdoctoral fellowship.

CORRESPONDENCE should be addressed to F.W. (e-mail: wudl@physics.ucsb.edu).

Site-ordering effects on element partitioning during rapid solidification of alloys

H. Assadi & A. L. Greer

Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK

WHEN an alloy solidifies, the component elements are often redistributed between the solid and liquid phases, so that the composition of the growing solid differs from that of the liquid. This effect, known as solute partitioning¹, is the main cause of inhomogeneity in alloys, affecting both the nature and the mechanical properties (for example, brittleness) of the solidified material. If solidification is rapid, the solid–liquid interface is no longer in equilibrium, and it is generally accepted—and has been demonstrated² for solid solutions—that this gives rise to reduced partitioning, in the limit leading to a solid having the same composition as the liquid through a process known as ‘solute trapping’³. We have recently argued⁴ on theoretical grounds that, when the solidifying phase shows site ordering, the partitioning behaviour can be considerably more complex: rapid solidification might lead to increased partitioning, a change in the direction of partitioning, or an absence of partitioning at solidification rates much lower than expected. Here we report the experimental verification of this phenomenon by demonstrating inverted partitioning during rapid solidification of the intermetallic compound NiAl.

We explore the partitioning behaviour in the solidification of intermetallic compounds by examining a model binary system in which the solid exhibits an order–disorder transformation of the

second order. The liquid is described as a regular solution and the solid is a body-centred cubic (b.c.c.) phase with long-range order parameter η (Fig. 1), described as a Bragg–Williams phase; the thermodynamic properties of the disordered solid are obtained by taking $\eta = 0$. By reasonable choice of thermodynamic parameters (details in ref. 4), it is possible to construct a model system in which the ordered solid phase and its thermodynamically unstable disordered version, if solidified, would show elemental partitioning in opposite directions. This is obtained, for instance, in a system in which the intermetallic phase, formed of elements with melting temperatures of 500 K and 1,000 K, shows a second-order transition at the critical temperature of 2,500 K and a congruent melting temperature of 1,100 K (that is, the highest melting temperature of the ordered solid, at which the liquid freezes with the same composition as the solid). As shown in the calculated phase diagram (Fig. 2), partitioning during the solidification of the stable ordered phase (heavy solid lines, corresponding to η at equilibrium) and the unstable disordered phase (dotted lines, corresponding to $\eta = 0$) would be in the same direction for $x_B < 0.5$ and in the opposite directions for $x_B > 0.5$ (where x_B is the mole fraction of B. In the latter case the partitioning should invert if the long-range order parameter were depressed, and hence a disordered rather than ordered solid is formed).

It has been shown theoretically⁵ and verified experimentally⁶ that increased solidification rates can result in the depression of the long-range order parameter. Boettinger and Aziz⁷ extended the theory of ‘solute trapping’⁷ to systems that may be decomposed into two sublattices (of the kind shown in Fig. 1), using a set of three response functions which describe the relationships between the growth velocity V and the interface variables (liquid and solid compositions, temperature and order parameter). The first function is the relation between the growth velocity and the change in free energy on solidification (ΔG): $V = V_c \{1 - \exp[\Delta G/(RT^*)]\}$, in which V_c is a kinetic rate parameter, R is the gas constant and T^* is the temperature of the interface. The other two functions are obtained by considering the diffusive-type reactions between the liquid and each sublattice of the solid. The fluxes of atoms across the interface are written as the rates of a pair of reactions, representing the exchange of atoms from their initial states on the α or β sublattice sites, over an energy barrier to their final states in the liquid⁵:

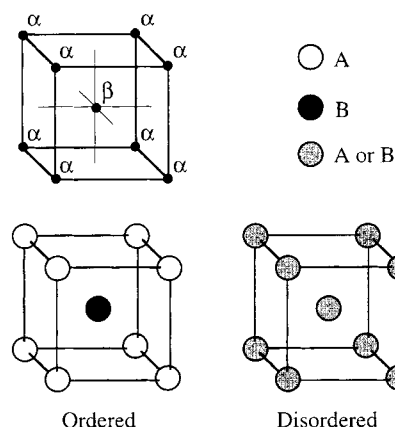
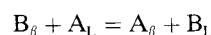
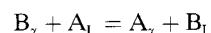


FIG. 1 The crystal structure of ordered and disordered solids in the model system. The top figure shows the α and β sublattice sites in a B2 superlattice. The long-range order parameter η (defined as the difference of the sublattice compositions) is zero for the disordered, and one for the perfectly ordered, solid. The equilibrium degree of order of the solid at a given temperature and composition is determined by minimizing the free energy of the solid with respect to η .

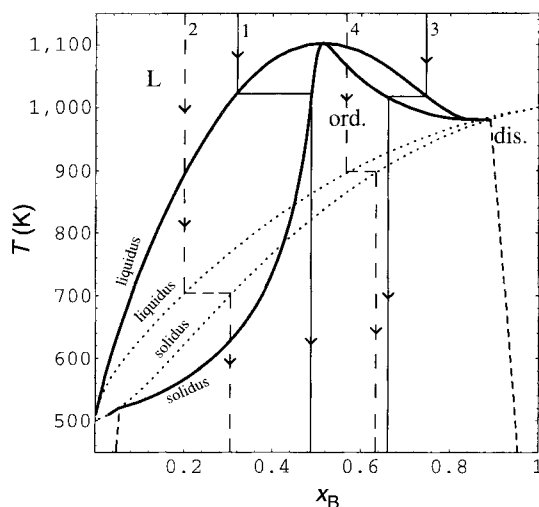


FIG. 2 Phase diagram of the model system. Solid lines show the stable equilibrium between liquid and the ordered solid, dotted lines the unstable equilibrium between liquid and the (hypothetical) disordered solid. The horizontal displacements of the vertical lines indicate, schematically, the directions of composition change brought about by solute partitioning at the solid-liquid interface. For mole fraction of solute $x_B < 0.5$, the partitioning is in the same direction for the ordered and disordered phase (lines 1 and 2). For $x_B > 0.5$, the partitioning is in opposite directions in the two cases (lines 3 and 4), a qualitative result that can easily be tested. The liquid is described as a regular solution with the interaction parameter $\Omega_L = -34,911 \text{ J mol}^{-1}$ and the solid as a Bragg-Williams phase with the nearest-neighbour interaction parameter $\Omega_S = -41,570 \text{ J mol}^{-1}$. The free-energy difference between solid and liquid for the pure elements is taken as $2R(T - T_{A \text{ or } B})$, in which R is the gas constant and $T_{A \text{ or } B}$ is the melting temperature of A or B. The dashed line is the locus of the order/disorder transition temperatures.

These fluxes of atoms are then linked to the growth velocity by conserving mass. The fluxes can be expressed in terms of variables such as solid and liquid compositions, the order parameter and the temperature at the solid-liquid interface; thus the conservation of mass is a basis for the two response functions linking the interface variables to the growth velocity. Applying these functions to model systems⁵ gives a transition, 'disorder trapping', in which η falls to zero, discontinuously or continuously (first- or second-order transitions, respectively) at a critical velocity, beyond which the growth of solid with $\eta \neq 0$ (and hence some degree of order) is kinetically impossible.

To analyse the effects on partitioning of the depression of the order parameter, we apply the theory of 'disorder trapping'⁵ to the model system (Fig. 2). Simultaneous solution of the corresponding interface response functions yields relationships between pairs of variables at the interface when one of the others is set constant. Figure 3 shows the relationship between liquid composition x_B^L , and a normalized growth velocity $v_n = V/V_D$ (where $V_D = D_i/\lambda$ is the atomic diffusivity speed in which D_i is the interface diffusivity and λ is a jump distance) for constant solid composition x_B^S . For $x_B < 0.5$, Fig. 3a, the partitioning during the growth of the ordered phase decreases with increasing velocity, as generally expected. Beyond the disorder trapping point (marked D) the behaviour is described by the 'solute trapping' model for a disordered solid solution. For the composition of the congruent freezing point, however, the partitioning increases with increasing v_n until the disorder trapping point is reached (Fig. 3b). As expected, the most interesting behaviour is found (Fig. 3c) for $x_B > 0.5$. For small v_n , partitioning occurs by rejection of B into liquid (as shown by the equilibrium solidus and liquidus on Fig. 2). As v_n is increased, the partitioning inverts at a point (marked I) where there is zero partitioning. After inversion, partitioning increases (by rejection of A

into liquid) and finally decreases asymptotically to zero.

Even without a quantitative kinetic model, it can be predicted that in this case the partitioning behaviour of the ordered phase always approaches that of the disordered phase as the growth velocity is increased; this must be so, as the ordered phase in a system showing a second-order transition (both thermodynamically and kinetically) increasingly resembles the disordered phase as the order parameter is depressed. It should be noted that this is generally not the case for the first-order transitions, in which the variation of thermodynamic functions with order parameter may not be monotonic.

In its basic features the model system (Fig. 2) resembles the metastable phase equilibria in the NiAl system involving the liquid, the ordered compound around the equiatomic composition (NiAl, β -phase, B2 structure, 'ordered' in Fig. 1) and the metastable (or unstable) disordered form of the compound (b.c.c. solid solution, A2 structure, 'disordered' in Fig. 1). The A2 to B2 ordering is expected to be second order, as in the model system. The similarity between the model system (taking component B to be Ni) and the A2-B2-liquid equilibria is sufficient to permit NiAl to be used qualitatively to test the partitioning predictions in Fig. 3. In particular, according to the predictions of Fig. 3b rapid solidification should result not in the congruent freezing of stoichiometric NiAl, but in the partitioning of aluminium into the liquid. Moreover, the prediction of Fig. 3c suggests that the partitioning in alloys of nickel-rich NiAl should invert, from the rejection of nickel to the rejection of aluminium into the liquid, as the growth velocity is increased. This has been tested by pulsed-laser melting of the surface of samples of nominal compositions of 35 and 50 at.% Al. The samples were cylinders (7 mm in diameter and 5 mm long) with ground faces to enhance the absorption of the laser beam. A 400-W laser (of wavelength 1,064 nm) with a pulse duration of 10 ms gave a molten zone ~ 1 mm in diameter and a few hundred micrometres deep. By calculation, this would give a resolidification velocity of the order of 0.1 m s^{-1} . After the

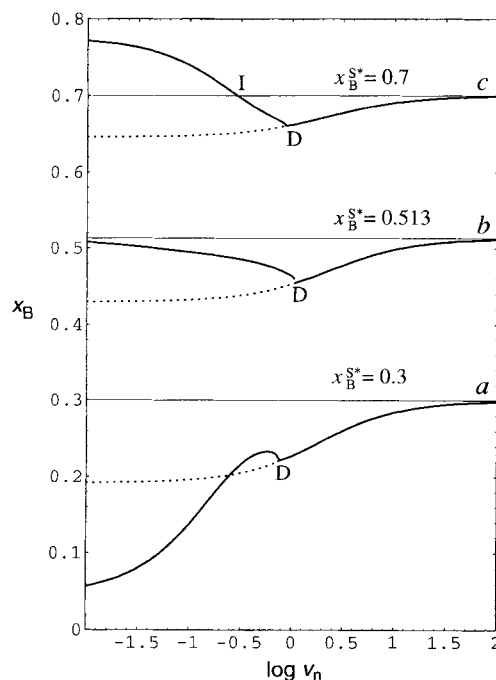


FIG. 3 The liquid composition at the interface as a function of growth velocity, calculated for three fixed solid compositions in the system shown in Fig. 2: a, $x_B^S = 0.3$; b, $x_B^S = 0.513$, congruent point; c, $x_B^S = 0.7$. The dotted lines correspond to unstable disordered solid. Disorder trapping is at point D. In each case complete solute trapping ($x_B^S = x_B^L$) occurs at high velocity. However, partitionless growth, point I in c, can occur at a velocity lower than that for disorder.

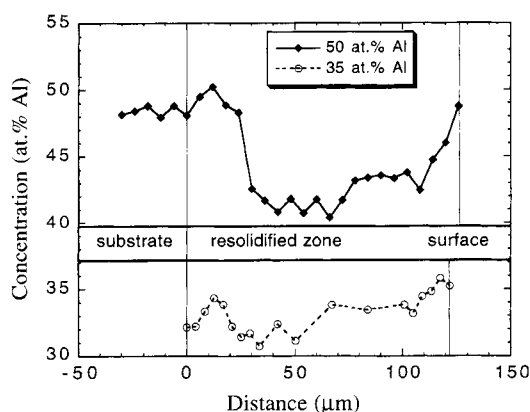


FIG. 4 The composition profiles (measured by EDS) along a trace on the cross-section of the resolidified zone in the pulsed-laser-treated samples with nominal compositions of 35 and 50 at.% Al. The decrease in the average aluminium content of the resolidified zone is an indication of aluminium loss, possibly due to evaporation. However, the rise of aluminium concentration near the surface of both samples indicates solute partitioning by rejection of aluminium into the liquid at the later stages of regrowth.

laser treatments, the resolidified zones were sliced perpendicular to the treated surface. The cross-sections of resolidified zones were initially examined by optical metallography; no dendritic structure was observed. Energy-dispersive spectrometry (EDS) along a trace on the sections (Fig. 4) confirms that there is significant compositional inhomogeneity in both samples as a result of rapid solidification. The decrease of the average aluminium content of the resolidified zones (especially in the NiAl sample with 50 at.% Al) indicates the possible loss of aluminium due to evaporation. Both profiles show a local maximum in the resolidified region near the substrate. This initial rise in the aluminium content can be interpreted as follows. The initial growth velocity in the region near the substrate, where non-equilibrium effects are negligible, is expected to be relatively low. Therefore, according to the equilibrium phase diagram (Fig. 2), the initial solid growing into a nickel-rich (B-rich) liquid would contain more Al than the liquid. However, as solidification proceeds, the interface accelerates and the non-equilibrium effects emerge. The solid grows into a liquid which is even more enriched in Ni, owing to the evaporation of Al. These two factors may explain the subsequent drop in the Al concentration. But the most important result is the rise in the Al content of the solid near the surface (the right of both composition profiles) indicating the rejection of Al into the liquid during regrowth; this could happen only if the partitioning had inverted (as in Fig. 3c). As shown by the two composition profiles, overall loss of Al does not affect the (predicted or observed) direction of the partitioning. Indirectly, the experiments provide evidence that rapid solidification can induce chemical disorder even in NiAl, for which the disordered state cannot readily be retained⁸. The disordered state persists only in a transient layer behind the solid-liquid interface, but has profound effects on partitioning.

These findings emphasise the differences in the solidification behaviour between ordered and disordered phases due to the energetic effects of site-ordering. In particular, there is the unprecedented demonstration that rapid growth may result not in reduced partitioning—always regarded as an advantage of rapid solidification processes—but in increased partitioning. Such an increase in partitioning may lead to the appearance of second phases and is likely to be a problem in welding and surface treatments of intermetallic compounds. On the other hand, the existence of a zero-partitioning point (for example I in Fig. 3) may be of use in producing chemically uniform intermetallic compounds where significant segregation would normally be expected.

A detailed kinetic analysis of the solidification of the ordered phase in this model system, or of γ' phase in the NiAl system⁹, shows that the inversion of partitioning can, for certain values of growth velocity and composition, be associated with the liquidus temperature falling below that of the solidus; in this case there is 'solutal superheating' (as opposed to 'solutal undercooling' which makes the interface morphologically unstable¹⁰, and absolute stability of the solid-liquid interface in a way essentially different from that due to capillarity effects¹¹. Again, this is of practical relevance as planar-front solidification is often desirable in the directional solidification of intermetallic compounds, but may be difficult to achieve in these systems with high thermal conductivity.

It follows from modelling and experiment that the reliable analysis of solidification in systems in which the major constituent is a compound will be possible only if account is taken of the chemical-order effects. □

Received 7 March; accepted 23 July 1996.

1. Chalmers, B. *Principles of Solidification* 126 (Wiley, New York, 1964).
2. Smith, P. M. & Aziz, M. J. *Acta Metall. Mater.* **42**, 3515–3525 (1994).
3. Kar, A. & Mazumder, J. *Acta Metall. Mater.* **40**, 1873–1881 (1992).
4. Assadi, H. & Greer, A. L. *Mater. Res. Soc. Symp. Proc.* (in the press).
5. Boettinger, W. J. & Aziz, M. J. *Acta Metall.* **37**, 3379–3391 (1989).
6. West, J. A., Manos, J. T. & Aziz, M. J. *Mater. Res. Soc. Symp. Proc.* **213**, 859–864 (1991).
7. Aziz, M. J. *J. Appl. Phys.* **53**, 1158–1168 (1982).
8. Koch, C. C. *Int. Mater. Rev.* **33**, 201–219 (1988).
9. Assadi, H. & Greer, A. L. *Iron Steel Inst. Jpn Int.* **35**, 574–579 (1995).
10. Tiller, W. A., Jackson, K. A., Rutter, J. W. & Chalmers, B. *Acta Metall.* **1**, 428–437 (1953).
11. Mullins, W. W. & Sekerka, R. F. *J. Appl. Phys.* **34**, 444–451 (1964).

ACKNOWLEDGEMENTS. We thank R. W. Cahn and J. M. Robinson for provision of samples and experimental assistance.

CORRESPONDENCE should be addressed to A.L.G. (e-mail: alg13@cam.ac.uk).

Recent changes in tropical freezing heights and the role of sea surface temperature

Henry F. Diaz* & Nicholas E. Graham†

* NOAA/ERL/CDC, 325 Broadway, Boulder, Colorado 80303, USA

† Climate Research Division, Scripps Institution of Oceanography, La Jolla, California 92093, USA

A WIDESPREAD retreat of alpine glaciers¹ and melting of tropical ice-cap margins^{2–7} has been observed in recent decades, over which time a general climate warming at lower altitudes has been documented⁸. Moreover, some ice-core records provide evidence suggesting that mid-tropospheric temperatures in the tropics have been greater in recent decades than at any time during the past 2,000–3,000 years⁷. Here we examine the processes controlling mountain glacier retreat by comparing high-altitude air-temperature measurements for the past few decades, to the temperatures predicted by a model atmosphere forced by the observed global pattern of sea surface temperature in a 19-year simulation⁹. The comparison strongly indicates that the observed changes in freezing-level height (the altitude of the 0 °C isotherm) are related to a long-term (over decades) increase in sea surface temperature in the tropics, and the consequent enhancement of the tropical hydrological cycle. Although changes in this cycle are likely to affect high-elevation hydrological and ecological balances worldwide^{10,11}, tropical environments may be particularly sensitive because the changes in tropical sea surface temperature and humidity may be largest and most systematic at low latitudes.

Some recent studies of high-elevation records in Europe^{12,13} have shown that surface air temperature measured at several isolated mountain peaks has risen by more than 1 °C during the

past century. The inferred warming is consistent with oxygen isotopic changes ($\delta^{18}\text{O}$) which indicate warming in the past 100 years, with the greatest change occurring in the past few decades⁷. Although such isotopic records are open to various interpretations, the implied temperature record is largely consistent with similar records constructed from sea surface temperatures (SST) and air-temperature records on land^{8,14}. Furthermore, there is clear evidence¹⁵ from at least one tropical glacier that extensive surface melting occurred sometime during the past two decades, for the first time in several hundred years. Other studies^{16–18} have documented upward trends in tropical tropospheric temperature and moisture content in recent decades^{16–18}.

Here we compare the observed changes in freezing-level heights, derived from radiosonde measurements in the tropics, to those obtained using the results of a numerical model simulation of the atmospheric response to observed SST in the period 1970–88. Further comparison of the temperature changes in high-elevation environments is made by using 40–45 years of surface temperature records from weather stations situated above 1,000 m in elevation, and vertical temperature profiles from about 75 radiosonde stations^{18,19} in the tropics (the region from 30° N to 30° S).

The numerical simulations were performed with the ECHAM2 and ECHAM3 climate models (Max Planck Institute) configured to give spatial resolution of approximately 5.5° and 2.8° latitude (T21 and T42 spectral truncations). Nineteen-year climate simulations, with prescribed global observed SST for the period 1970–88, were conducted with each model⁹. Two sets of radiosonde data were used to evaluate the modelling results: a set of 65 radiosonde stations distributed throughout the tropics from about 30° N to 30° S, with data available for the period 1970–86, and a second set of 10 stations in South America¹⁹, with a longer record (1958–90) which were situated mainly in a meridional direction. The air-temperature observations from the 12:00 GMT daily soundings (except for Lima in Peru, where only the 00:00 GMT soundings were available) were used to estimate the geopotential height of the local freezing surface by simple linear interpolation between the two levels straddling the 0 °C air-temperature value. The data contained information at significant level heights, as well as at the standard pressure levels. The monthly mean height of the 0 °C surface was calculated by averaging the calculated, daily freezing-level heights.

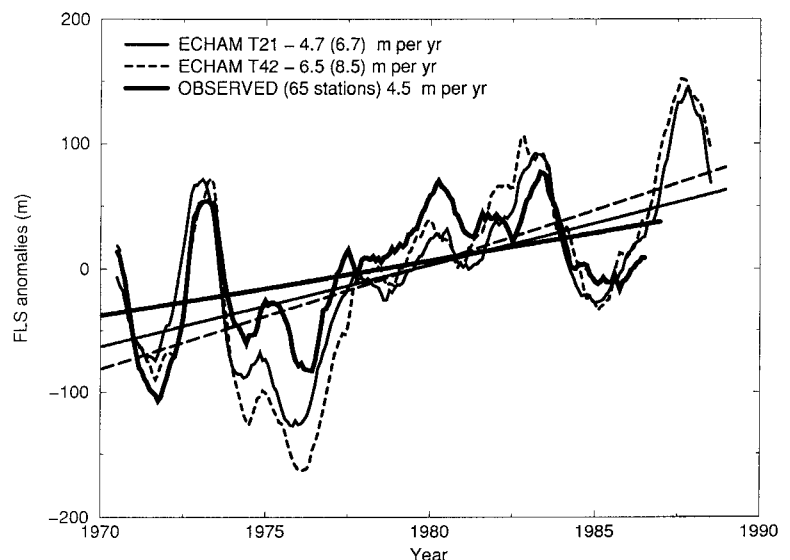
Figure 1 shows the evolution of the freezing-level surface (FLS) anomalies (in each instance, the mean over the period of record for each calendar month was removed) averaged over the tropics (30° N–30° S) in the T21 and T42 model runs (the heavy solid and heavy dashed curves, respectively), together with the correspond-

ing value (thin solid line) obtained by averaging the values from the data from the 65 radiosonde stations binned into 5° latitude bands. The temporal increase of the freezing-level height from 1970 to the late 1980s, in the deep tropics (~15° N–15° S), amounts to about 110 m (calculated from the total linear trend over the analysis period). A latitude profile comparison of the simulated annual-mean freezing-level height changes, for the two models, is given in Fig. 2. For comparison, the equivalent changes based on the radiosonde data are shown by large filled circles. The patterns are quite similar, but the magnitude of the changes in the T42 run is greater, and both simulations may underestimate the observed changes in FLS in the deep tropics (~10° N–10° S). There are substantial differences between the model results and the observations, at latitudes less than about 10° N, which reflect, at least in part, sampling differences between them. There are also seasonal differences (not shown) in FLS height which are evident in the simulation, but the critical feature is the general increase of FLS height in the tropics during this time. The spatial detail of the T42 simulation is more realistic and is the one considered below.

Figure 3 illustrates the spatial distribution of annual mean FLS changes in the T42 model simulation, which indicates that the pattern of freezing-level height changes in the model strongly resembles the upper-tropospheric anomaly pattern which occurs during the course of El Niño events in the tropical Pacific Ocean^{20,21}. Some of the greatest simulated warming occurs in the region of Peru where both the Quelccaya and Huascarán ice caps^{2,7} are located (see Fig. 3). The question is whether this mid-tropospheric warming is associated with the recent ~20-yr episode of generally above average SST in the tropical Pacific^{9,22–24}. One of the studies⁹ of this episode, a synthesis of observational and model data, provides evidence indicating that the increases in tropical SST during this period result in an enhancement of the tropical hydrological cycle. One manifestation of such an enhancement is a trend towards larger temperature increases with height in the tropical troposphere. This is a consequence of a tendency for the vertical temperature structure to approach more closely a moist, adiabatic lapse-rate.

A further test of the SST–FLS link is shown in Fig. 4. The annual mean changes in FLS, for the set of 10 long-term radiosonde stations in South America, are depicted here in the form of box plots which represent the distribution of anomalous FLS height for each year, among the 10 South American radiosonde stations. There is a notable jump in FLS heights which occurs around 1976–77. Because tropical SST changes in the 1970s have been implicated in a variety of anomalous geophysical events²⁴, we compare the observed changes in tropical SST, from the late 1950s to 1990, to the corresponding changes in FLS in South America

FIG. 1 Anomaly time series of the average height (in metres) of the 0 °C surface (the FLS) in the tropics (1970–88), derived from the ECHAM2–T21 and ECHAM2–T42 simulation as described in the text. Also shown is the equivalent change in freezing-level height from a network of 65 radiosonde stations located from 30° N to 30° S for the period 1970–86. Linear trends of 4.7 and 6.5 m yr⁻¹ for FLS height in the T21 and T42 simulations are for the period 1970–86. The values in parentheses (6.7 and 8.5 m yr⁻¹ for T21 and T42, respectively) are for 1970–88.



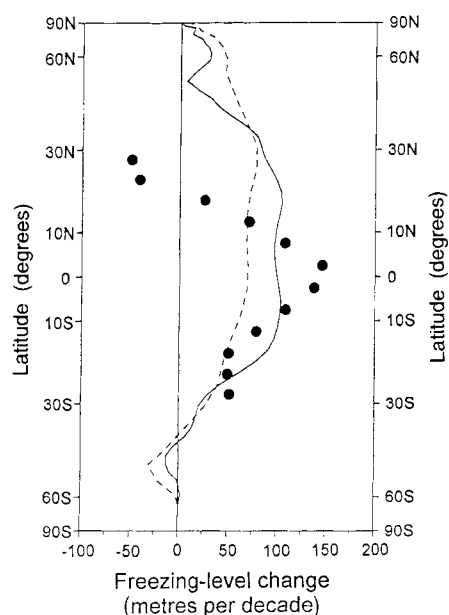
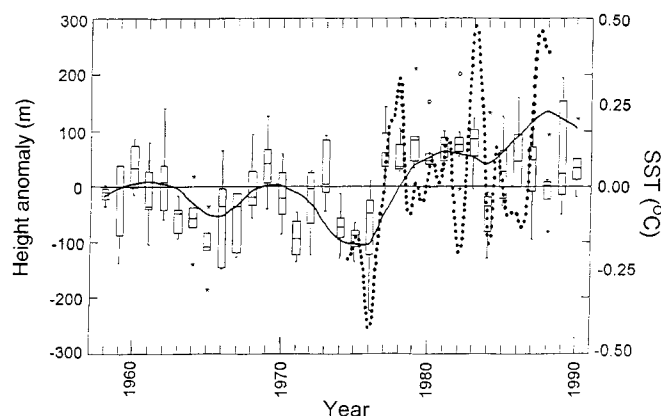
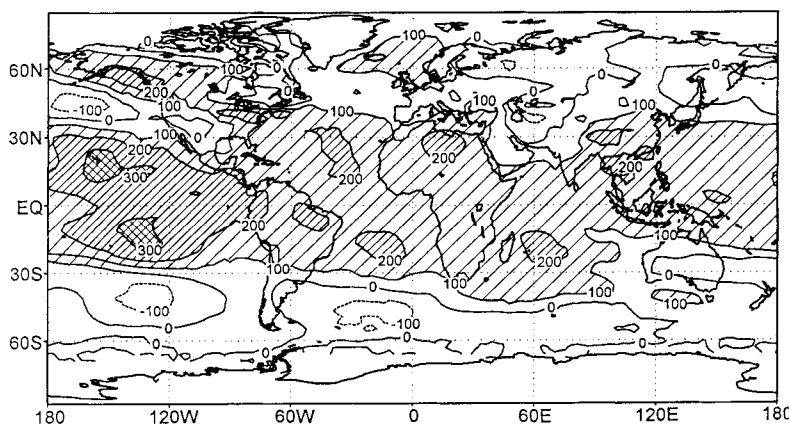


FIG. 2 A comparison of the total linear trends (in metres) of the annual mean freezing-level heights, in an 18-yr global climate model (GCM) simulation with the T21 (dashed line) and T42 ECHAM2 models, as a function of latitude, with the corresponding changes in the observations (solid dots). The linear trend changes for the radiosonde data (1970–86) were binned into 5° boxes and the median value within each 5° band was plotted.

FIG. 3 A comparison of the spatial distribution of the total linear trends (in metres) in annual mean freezing-level heights in an 18-yr GCM simulation with the ECHAM3–T42 model version.



(Fig. 4, smooth solid line). To underline the connection between the changes in the hydrological cycle in the Pacific region and the SST and FLS height, an index of the precipitation changes in the tropical Pacific²⁵ is also shown (smooth dotted line). The results in Fig. 4 suggest a fairly tight coupling of the tropical SST (arising mostly from changes in the tropical Pacific Ocean), rainfall changes in the Pacific and the freezing-level heights over South America. The rainfall variations are large and strongly influenced by the El Niño/Southern Oscillation phenomenon²⁵. However, the relative changes between all the variables are consistent over time.

One further observational test was performed to address the latitudinal representativeness of the changes found in the radiosonde data. We have calculated linear (surface) temperature trends for a longer period of temperature observations (1951–94) using all available data from stations situated more than 1,000 m above sea level. The station data are part of the US National Climatic Data Center archives and were binned into 20°-bands from 30° S to 30° N, with a fourth band made up of all available stations located at latitudes greater than 30° N. The resulting temperature trends are consistent with the changes in FLS height derived from the set of radiosonde observations described above, and the modelling experiments (Fig. 2). That is, surface air-temperatures in higher elevation (>1,000 m) tropical regions have risen by a few tenths of a degree Celsius since about 1950, and there is some indication that the changes are greater in magnitude than the corresponding changes in the extratropics, which are only marginally positive during this period.

Our analysis provides evidence for a rise in tropical freezing-level surface elevation that is closely coupled to increases in tropical SST during the past few decades. This change is consistent with the observations of the melting of high-elevation tropical ice

FIG. 4 Distribution of annual freezing-level height anomalies for 10 radiosonde stations in South America for the period 1958–90. The results are presented in terms of 'box and whisker' plots, which show approximately the 5th, 25th, 50th and 95th percentiles of the FLS height anomaly with respect to the station means, and values falling outside these limits, indicated by asterisks or open circles (99th percentile). The stations used were: Bogota, Colombia; Manaus, Brazil; Lima, Peru; La Paz, Bolivia; Antofagasta, Quintero, Puerto Montt and Punta Arenas, Chile; Salta and Rio Grande, Argentina. The smoothed solid line represents the low-pass filtered, median tropical SST anomalies (seasonal values referenced to the 1951–92 period) based on a 5° gridded dataset, and the solid dots correspond to the time coefficients (arbitrary units, scaled to the SST axis) of the first principal component of tropical Pacific annual mean rainfall²⁵.

in the Andes. Although changes in temperature are probably the dominant mechanism which produces this melting, there is also evidence, from both observations and models, that the changes in temperature have been accompanied by increases in moisture content in the lower troposphere^{17,18}, which is consistent with the increase of the tropical SST and the resulting enhancement of the tropical hydrological cycle. These increases in humidity would also contribute to accelerated melting rates⁴.

On the basis of very long isotopic records obtained from the various tropical ice-cap cores (of the order of several centuries to

several thousand years), the warmth recorded in the tropical oceans in several past decades may perhaps be at an unprecedented level since the mid-Holocene period (~3,000–4,000 yr BP)⁷. Whether this recent increase is natural, anthropogenic, or both, remains an open question. Recent evidence^{1,11}, however, suggests that high-elevation environments may be particularly sensitive to long-term changes in tropical SST and tropospheric humidity, which are likely to have an impact on the hydrological and ecological balances of high-altitude zones throughout the globe, perhaps to the greatest extent in the tropics. □

Received 6 March; accepted 23 July 1996.

- Oerlemans, J. *Science* **264**, 243–245 (1994).
- Thompson, L. G. et al. *Global Planet. Change* **7**, 145–156 (1993).
- Brecher, H. H. & Thompson, L. G. *Photogramm. Eng. Remote Sensing* **59**, 1017–1022 (1993).
- Hastenrath, S. & Kruss, P. *Ann. Glaciol.* **16**, 127–133 (1992).
- Kaser, G. & Noggler, B. *J. Glaciol.* **37**, 315–318.
- Schubert, C. *Erdkunde* **46**, 58–64 (1992).
- Thompson, L. G. et al. *Science* **269**, 46–50 (1995).
- Intergovernmental Panel on Climate Change (IPCC) *Climate Change '95, The Science of Climate Change* (Cambridge Univ. Press, 1996).
- Graham, N. E. *Science* **267**, 666–671 (1995).
- Schneider, S. H. *Bull. Am. Meteorol. Soc.* **71**, 1292–1304 (1990).
- Beniston, M. (ed.) *Mountain Environments in Changing Climates* (Routledge, London, 1994).
- Beniston, M., Rebetez, M., Giorgi, F. & Marinucci, R. *Theor. Appl. Climatol.* **49**, 135–159 (1994).
- Weber, R. O., Talkner, P. & Stefanski, G. *Geophys. Res. Lett.* **21**, 673–676 (1994).
- Jones, P. D., Bradley, R. S. & Jouzel, J. (eds) *Climatic Variations and Forcing Mechanisms of the Last 2000 Years* (NATO ASI Series, Springer, Berlin, 1996).
- Thompson, L. G. *Melting of our Tropical Climatic Archives* (PAGES Core Project Office, Bern, Switzerland, 1995).

- Hense, A. P., Krahe, P. & Flohn, H. *Meteorol. Atmos. Phys.* **38**, 215–227 (1988).
- Gutzler, D. S. *Geophys. Res. Lett.* **19**, 1595–1598 (1992).
- Gaffen, D. J., Barnett, T. P. & Elliott, W. P. *J. Clim.* **4**, 989–1008 (1991).
- Eskridge, R. E. et al. *Bull. Am. Meteorol. Soc.* **76**, 1759–1775 (1995).
- Newell, R. E. & Wu, Z.-X. *J. Geophys. Res.* **97**, 3693–3709 (1992).
- Yulaeva, E. & Wallace, J. M. *J. Clim.* **7**, 1719–1736 (1994).
- Trenberth, K. E. *Bull. Am. Meteorol. Soc.* **71**, 998–993 (1991).
- Graham, N. E. *Clim. Dyn.* **10**, 135–162 (1994).
- Miller, A. J., Cayan, D. R., Barnett, T. P., Graham, N. E. & Oberhuber, J. M. *Oceanogr.* **7**, 21–26 (1994).
- Morrissey, M. L. & Graham, N. E. *Bull. Am. Meteorol. Soc.* **77**, 1207–1219 (1996).

ACKNOWLEDGEMENTS. This work was partially supported by the US Department of Energy and the US National Oceanic and Atmospheric Administration. It grew out of discussions at a meeting in Wengen, Switzerland in September 1995 which was funded by several US and European agencies and organized by H.F.D. and M. Beniston of the Swiss Institute of Technology (ETH) in Zürich. We thank M. Beniston, R. Bradley, S. Hastenrath, J. Hurrell, U. Schotterer and L. Thompson for discussion, L. Bengtsson and E. Roeckner for providing the results from the MPI models, and T. Barnett and D. Gaffen for providing the 65 station radiosonde dataset.

CORRESPONDENCE should be addressed to H.F.D. (e-mail: hfd@cdc.noaa.gov).

Microbial generation of economic accumulations of methane within a shallow organic-rich shale

Anna M. Martini*, Joyce M. Budai*, Lynn M. Walter* & Martin Schoell†

* Department of Geological Sciences, University of Michigan, 2534 C C Little Building, 425 East University, Ann Arbor, Michigan 48109-1063, USA

† Chevron Petroleum Technology Company, 1300 Beach Boulevard, La Habra, California 90631, USA

ALTHOUGH methane of bacterial origin is ubiquitous in marine and freshwater sediments, economic accumulations of bacterial gases occur mainly at depths of several kilometres in Tertiary basins that had high sedimentation rates^{1,2}. Here we present an integration of geochemical and isotopic data from gas and water extracted from the Upper Devonian Antrim shale, along the northern margin of the Michigan basin, which demonstrates that significant volumes of bacterial gas have been generated in organic-rich shales at depths of less than 600 metres. The Antrim shale is mainly a self-sourced reservoir, in contrast to conventional gas deposits that have migrated from a source to a reservoir, and has become one of the most actively exploited gas reservoirs³ in the United States. The gas-forming processes operating at shallow depths in the Antrim shale are not unique⁴, and an understanding of these processes should lead to the identification and development of other economic, non-conventional gas deposits around the world.

The Antrim shale is rich in organic matter (up to 20 wt%), comprised of hydrogen-rich algal material and is thermally immature (vitrinite reflectance $R_o = 0.4$ – 0.6) in the area of study⁵. Our study area is along the northern margin of the Michigan basin

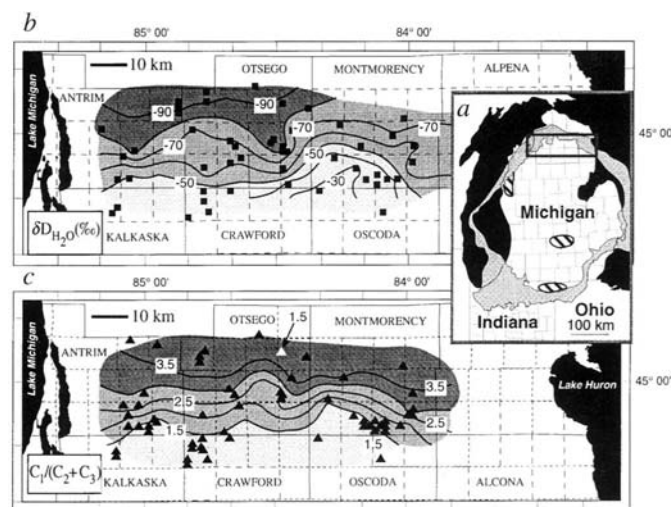


FIG. 1 a, Study area at the margin of the Michigan basin in northern Michigan, USA with the subcrop of the Antrim shale shown in grey. Over 5,000 wells have been completed in the shale for commercial production of methane. Of these, over 95% were commercially successful. Estimates³ of total Antrim shale reserves for the Michigan basin are in the range $(9-29) \times 10^{11} \text{ m}^3$. Gas and co-produced water were collected at the well-head. Sampling depths ranged from 200 m on the northern margin of the producing zone, to nearly 650 m in the south. Striped areas indicate regions of new exploration. b, Gradients in δD values of water (shown on contours) occur over the producing trend due to freshwater ($\delta D \approx -90\text{‰}$) influx mixing with basinal brines ($\delta D \approx -30\text{‰}$). Well locations are shown as filled squares. The δD value of the water was determined using a Finnigan MAT Delta S ratio mass spectrometer, with a precision of $\pm 2\text{‰}$. Waters were reacted with zinc metal to liberate hydrogen gas²⁸. c, A plot of $\log(C_1/(C_2 + C_3))$ for produced gas from the Antrim shale, showing a steep gas compositional gradient (with increasing C_2 -content) from the margin towards the centre of the basin. The value for one anomalous well (shown as an open triangle) is not contoured as contamination is suspected. Gas samples were obtained from the well-head using stainless-steel cylinders, and analysed by routine gas chromatographic techniques.

(Fig. 1a) where the Antrim shale subcrops under a variable thickness of Pleistocene glacial drift. This region experienced multiple glacial advances and retreats⁶ which greatly modified the hydrological and pressure regimes, possibly enhancing meteoric water recharge in the northern part of the basin⁷. Gas is produced from natural and enhanced fractures^{8,9} within the Antrim shale. Over 5,000 wells produce $1.2 \times 10^7 \text{ m}^3$ of gas per day (ref. 3), as well as large quantities of saline water. The gas is mainly composed of methane (CH_4), but also includes varying amounts of carbon dioxide (CO_2) and ethane (C_2H_6). Because most gas reserves are adsorbed within the organic shale matrix³, gas production rates are enhanced by lowering pressure at the borehole by means of dewatering.

The origin of gaseous hydrocarbons and carbon dioxide in the Antrim shale reservoir has been uncertain. An early study suggested that gas from the Antrim shale might be of microbial origin on the basis of its high methane content and shallow depths of production¹⁰. Recent research favours a thermogenic origin, based on $\delta^{13}\text{C}$ values of $\sim -50\text{‰}$ (w.r.t. the VPDB standard) for methane¹¹. However, the carbon isotope composition of methane by itself does not reveal its origin, as it varies directly with the $\delta^{13}\text{C}$ value of the carbon source. For example, bacterial methane generated in closed systems can have $\delta^{13}\text{C}$ values similar to that of incipient thermogenic methane (~ -60 to -40‰), depending on the $\delta^{13}\text{C}$ of the CO_2 which has been reduced by bacterial action¹²⁻¹⁵.

We examined the chemical and isotopic compositions of the gas and co-produced water from the Antrim shale. The relationship between the hydrogen isotope compositions of methane and co-produced water, in conjunction with the carbon isotope compositions of both carbon dioxide and methane¹⁶⁻¹⁸, suggest that the dominant source of methane in the Antrim shale is microbial methanogenesis⁹. Relatively recent water recharge induced strong isotopic and related chemical gradients over the producing trend, as illustrated by the δD values of water (Fig. 1b). Dilute ground water ($<0.1 \text{ M NaCl}$; δD more negative than -90‰) is recharged into the Antrim shale at the subcrop through an extensive fracture network^{8,9} and progressively mixes with basinal brine ($>4 \text{ M NaCl}$; δD more positive than -30‰) (Fig. 1b).

Carbon-14 data from dissolved inorganic carbon (DIC) in water from the Antrim shale constrain the recharge age to a maximum value of 22,000 yr BP (ref. 9). Concentrations of DIC in waters produced from the Antrim shale are far greater than those of the ground waters from glacial drift. These dilute waters enter the Antrim shale with approximately 3–6 mM DIC and become charged with DIC concentrations up to 70 mM. The DIC in waters from the Antrim shale has unusually positive $\delta^{13}\text{C}$ values of approximately $+30\text{‰}$ (w.r.t. VPDB)⁹ and ultimately is derived from the abundant immature organic matter present in the shale by means of microbial activity.

Direct evidence for a microbial metabolic pathway is the positive correlation of δD values of CH_4 and formation water produced from the same well. In environments similar to the Antrim shale, carbon dioxide reduction is the dominant microbial pathway^{4,14,18}. This pathway involves two steps. Bacteria metabolize organic compounds and water by means of enzymatic reactions to produce hydrogen and carbon dioxide^{16,19}. Experimental data suggests that the free hydrogen produced in this way exchanges with environmental water¹². Methanogens subsequently reduce the carbon dioxide using this hydrogen. Thus, the hydrogen isotope composition of methane produced by carbon dioxide reduction is dependent on the hydrogen isotope composition of the associated water (Fig. 2a, solid line). Gas and water from the Antrim shale have δD values which fall near the line with a slope of unity (Fig. 2a), suggesting that they are largely formed via the carbon dioxide reduction pathway. In contrast, fluids sampled from the stratigraphically lower carbonates of the Silurian Niagaran formation and in deeper-producing wells from the Antrim shale, nearer the centre of the basin, clearly deviate from the trend shown for microbial gases. These gases are of

thermogenic origin²⁰, and the hydrogen in methane has not exchanged with associated waters (Fig. 2a).

A microbial origin of methane from the Antrim shale is further supported by comparison to known bacterial gases from glacial drift deposits in the Illinois basin which share a similar carbon isotope fractionation between carbon dioxide and methane^{14,18}. In the Illinois basin, low $\delta^{13}\text{C}$ values of CH_4 occur, indicating an open system where the supply rate of CO_2 is rapid compared to the rate of methanogenesis. In a closed or CO_2 -limited system, CO_2 is progressively converted to methane and the $\delta^{13}\text{C}$ of residual CO_2 becomes increasingly more positive^{13,14,18}. The CO_2 in gases from

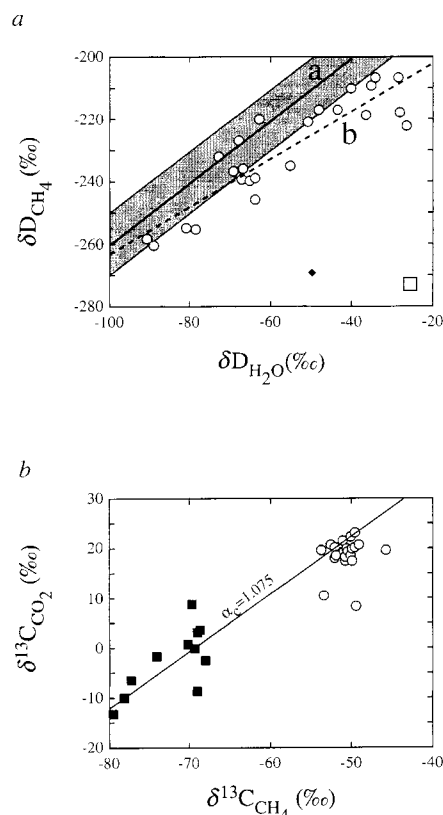


FIG. 2 a, Hydrogen isotopic composition of methane and associated (co-produced) water (open circles). The solid line 'a' follows the equation $\delta\text{D}_{\text{CH}_4} = \delta\text{D}_{\text{H}_2\text{O}} - (160 \pm 10)\text{‰}$ (refs 16, 19) with a slope of one (where $\delta\text{D}_{\text{CH}_4}$ is the δD value of CH_4 and $\delta\text{D}_{\text{H}_2\text{O}}$ is the δD value of H_2O) suggesting that all hydrogen in the methane is from water. This is consistent with microbial formation of CH_4 via CO_2 reduction ($\text{CO}_2 + 4\text{H}_2 \rightarrow 2\text{H}_2\text{O} + \text{CH}_4$). The area shaded in grey shows the variance ($\pm 10\text{‰}$) around the ideal line. The correlation of Antrim shale gas and water samples (dashed line 'b'; $y = 0.76x - 186.71$, $r^2 = 0.80$) is evidence of microbial methane production by CO_2 reduction. Mixing with thermogenic gases would move Antrim shale gas samples off the slope of one. Thermogenic gases form one Niagaran limestone location (filled diamond) and one Antrim shale mid-basin sample (open square) are outside the bacterial population. Carbon and hydrogen isotope analyses of gas were performed using VG 903 and VG 602 mass spectrometers, respectively²⁹. The δD values are given relative to the VSMOW standard, and the $\delta^{13}\text{C}$ values relative to the VPDB standard. Reproducibility for isotopic measurements are $\pm 1.7\text{‰}$ for δD and $\pm 0.1\text{‰}$ for $\delta^{13}\text{C}$. b, Plot of $\delta^{13}\text{C}$ values of CH_4 versus $\delta^{13}\text{C}$ of CO_2 for gases from the Antrim shale (open circles) and from gases produced microbially in the Illinois basin (filled squares)¹⁴. For reference, a fractionation factor (α_c) of 1.075 where $\alpha_c = (\delta^{13}\text{C}_{\text{CO}_2} + 10^3)/(\delta^{13}\text{C}_{\text{CH}_4} + 10^3)$ (suggestive of microbial processes) has been plotted¹⁵. Neither Silurian Niagaran formation gases nor the single Antrim shale gas sample from the central basin contain sufficient CO_2 for isotopic analysis.

the Antrim shale has $\delta^{13}\text{C}$ values up to +22‰ (VPDB). Carbon dioxide reducing methanogens metabolize this ^{13}C -enriched CO_2 and produce correspondingly positive $\delta^{13}\text{C}$ values of CH_4 (Fig. 2b).

The molecular composition of gases from the Antrim shale varies from pure methane to ~5 vol.% of hydrocarbons with two or more carbon atoms (C_2+). Virtually pure methane is produced near the basin margin, and gases with increasingly greater ($\text{C}_2 + \text{C}_3$) contents are produced deeper in the basin (Fig. 1c). These molecular ratio patterns are consistent with mixing, bacterial alteration or migration phenomena²¹. In general, higher-molecular-weight compounds such as ethane and propane tend to be retained in the rock, and lower-molecular-weight gases (that is methane) migrate more readily²². Mixing of two thermogenically sourced gases of differing thermal maturity could produce the $\text{C}_1/(\text{C}_2 + \text{C}_3)$ ratios observed, but should result in a shift of $\delta^{13}\text{C}$ values across the study area^{15,16}. The $\delta^{13}\text{C}$ values of CH_4 from the Antrim shale, however, show no coherent regional pattern (Fig. 3a). Diffusive migration can be ruled out because methane in the north has a different δD value from methane in the south, and fractionation of only one isotopic species is unlikely. Importantly, $\delta^{13}\text{C}$ values of ethane and propane (M.S. and A.M.M., unpublished data) increase with proximity to the subcrop. This pattern is consistent with bacterial oxidation, probably occurring during the initial influx of meteoric water at the subcrop. This process is not volumetrically significant as $\delta^{13}\text{C}$ values for methane do not increase at the subcrop of the Antrim shale²³ (Fig. 3a). Significantly, bacterial oxidation of a thermogenic gas would not produce the observed correlation between δD values for water and methane.

We suggest that the gas in the Antrim shale originates from the mixing of thermogenic gas (with $\text{C}_1/(\text{C}_2 + \text{C}_3) < 100$) and methane of microbial origin having similar $\delta^{13}\text{C}$ values of methane (Fig. 3b). Plausible sources for the thermogenic gas endmember are the Silurian Niagaran formation or the mid-basin Antrim Shale. Methane in gases from the Niagaran formation have $\delta^{13}\text{C}$ values of approximately -51‰ and $\text{C}_1/(\text{C}_2 + \text{C}_3)$ ratios from 2 to 10. Gas produced from the Antrim shale in the middle of the basin has a $\delta^{13}\text{C}$ value of -54‰, a $\text{C}_1/(\text{C}_2 + \text{C}_3)$ ratio of 9, and CO_2 content too low to measure. These findings suggest that both gases are thermogenic in origin^{15,16,24}. Thus, variations in the concentration of C_2+ hydrocarbons near the subcrop occur owing to a combination of bacterial oxidation and mixing of thermogenic and microbial gas (Fig. 1c).

Extensive microbial methanogenesis at shallow depth has important implications for the timing of water recharge and methane generation in the Antrim shale. The δD values of methane and co-produced water from the Antrim shale require that the waters have a residence time greater than that needed to generate the significant methane accumulation. This microbial gas generation in the Antrim shale occurred during or following the final stages of Pleistocene glaciation in northern Michigan based on ^{14}C ages of DIC in waters from the Antrim shale^{6,9}. The presence of ethane and propane represents the mixing of a relatively small fraction (up to 20%) of thermogenic gas with bacterial gas.

The connection among hydrology, geochemistry and gas production for the Antrim shale is not unique and will probably be applicable to other hydrogeological settings. Recent gas exploration and development in the Illinois basin targets the shallow margins of the Devonian New Albany shale, where a similar set of hydrogeological conditions exist. Non-conventional natural gas deposits, such as those in the Antrim shale, can be economic and provide ideal local energy resources because of shallow drilling depths and the geographically predictable nature of *in situ* production.

Widespread discoveries of microbial communities in the deep subsurface have led to speculation that biogenic gas may comprise a larger component of natural gas reserves than previously thought^{25,26}. Non-conventional gas resources, such as coal-beds

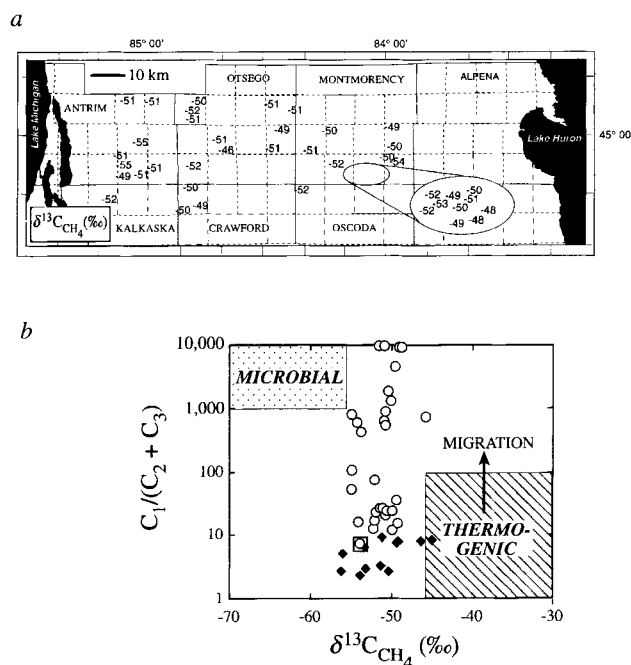


FIG. 3 a, $\delta^{13}\text{C}$ (w.r.t. VPDB) values for methane are plotted over well locations and are invariant across the trend. b, A plot of (methane/(ethane + propane)) versus $\delta^{13}\text{C}$ of methane for Antrim shale gases (open circles). Gases produced from the Silurian Niagaran limestone (filled diamonds) and Antrim shale (open square) samples from the central portion of the basin are thought to represent thermogenic gas sources²⁰. Carbon isotopic composition of methane has a narrow range from -47 to -56‰ (VPDB) despite three orders of magnitude variation in log (methane/(ethane + propane)). Fields for biogenic and thermogenic gases are included (adapted from ref. 23); values for Antrim shale gas samples are located at the edge of both fields on the plot. An arrow representing expected changes in gas composition due to migration of a thermogenic gas is also displayed.

and organic-rich shales have largely been attributed to thermogenic processes^{4,27}, yet they may contain far more microbial gas than previously believed. Our findings should stimulate a broader view of these resources and establish integrated geochemical approaches for recognition of biogenic gas deposits. □

Received 5 January; accepted 16 July 1996.

- Mattavelli, L., Ricchiuto, T., Grignani, D. & Schoell, M. *Bull. Am. Assoc. Petrol. Geol.* **67**, 2239–2254 (1983).
- Rice, D. D. *Gulf Coast Assoc. Geol. Soc. Trans.* **30**, 203–213 (1980).
- Scott, A. R., Kaiser, W. R. & Ayers, W. B. *Bull. Am. Assoc. Petrol. Geol.* **78**, 1186–1209 (1994).
- Frantz, J. *GasTIPS* **2**, 5–11 (1995/96).
- Rullkötter, J., Marzi, R. & Meyers, P. A. in *Early Organic Evolution: Implications for Mineral and Energy Resources* (eds Schidlowski, M. et al.) 324–335 (Springer, Heidelberg, 1992).
- Farrand, W. R. & Eschman, D. F. *Michigan Academician* **7**, 31–56 (1974).
- Bahr, J. M., Moline, G. R. & Nadon, G. A. *Mem. Am. Assoc. Petrol. Geol.* **61**, 153–165 (1994).
- Apotria, T. G., Kaiser, C. J. & Cain, B. A. in *Rock Mechanics* (eds Nelson, P. P. and Lauback, S. E.) 809–816 (Balkema, Rotterdam, 1994).
- Walter, L. M. et al. *Hydrogeochemistry of the Antrim Shale, Northern Michigan Basin* (Annual Rep., Gas Res. Inst., Chicago, 1996).
- Vugrinovich, R. J. *Petrol. Geol.* **11**, 429–442 (1988).
- Dellapenna, T. M. thesis, Western Michigan Univ., Kalamazoo (1991).
- Balabane, M., Gallimov, E., Hermann, M. & Letolle, R. *Org. Geochem.* **11**, 115–119 (1987).
- Jenden, P. D., Newell, K. D., Kaplan, I. R. & Watney, W. L. *Chem. Geol.* **71**, 117–147 (1988).
- Coleman, D. D., Liu, C. L. & Riley, K. M. *Chem. Geol.* **71**, 23–40 (1988).
- Fuex, A. N. J. *Geochem. Explor.* **7**, 155–188 (1977).
- Schoell, M. *Geochim. Cosmochim. Acta* **44**, 649–661 (1980).
- Woltemate, I., Whitcar, M. J. & Schoell, M. *Limnol. Oceanogr.* **29**, 985–992 (1984).
- Whitcar, M. J., Faber, E. & Schoell, M. *Geochim. Cosmochim. Acta* **50**, 693–709 (1986).
- Nakai, N., Yoshida, Y. & Ando, N. *Chikyū Kagaku* **7**, 87–98 (1974).
- Vary, J. A., Elenbaas, J. R. & Johnson, M. A. in *Natural Gases of North America 1761–1797* (Amer. Assoc. Petroleum Geologists, Tulsa, Oklahoma, 1968).
- Schoell, M. J. *Geol. Soc. Lond.* **140**, 415–422 (1983).
- Pinzhof, A. A. & Huc, A. Y. *Chem. Geol.* **126**, 281–290 (1995).
- Coleman, D. D., Liu, C., Hackley, K. C. & Pelphrey, S. R. *Environ. Geosci.* **2**, 95–103 (1995).
- Bernard, B. B., Brooks, J. M. & Sackett, W. M. J. *Geophys. Res.* **83**, 4053–4061 (1978).
- Stetter, K. O. et al. *Nature* **365**, 743–745 (1993).

26. Stevens, T. O. & McKinley, J. P. *Science* **270**, 450–453 (1995).
 27. Rice, D. D. in *Hydrocarbons from Coal* (eds Law, B. E. & Rice, D. D.) 159–184 (Amer. Assoc. Petroleum Geologists, Tulsa, Oklahoma, 1993).
 28. Venneman, T. W. & O'Neill, J. R. *Chem. Geol.* **103**, 227–234 (1993).
 29. Schoell, M., Jenden, P. D., Beeunas, M. A., Coleman, D. D. *Gas Technology Symp.* 337–344 (Soc. Petroleum Engineers, Calgary, 1993).

ACKNOWLEDGEMENTS. We thank J. R. O'Neill, E. J. Essene, B. K. Wilkinson, A. Young, L. Price, T. C. W. Ku & R. T. Klein for their comments. This work was supported by the Gas Research Institute, the American Chemical Society Petroleum Research Fund, Shell Western Exploration and Production and Chevron Petroleum Technology Company.

CORRESPONDENCE should be addressed to A.M.M. (e-mail: amartini@umich.edu).

Magnetic information affects the stellar orientation of young bird migrants

Peter Weindler, Roswitha Wiltschko
& Wolfgang Wiltschko

Fachbereich Biologie der Universität Frankfurt a.M., Zoologie,
Siesmayerstrasse 70, D 60054 Frankfurt a.M., Germany

WHEN young birds leave on their first migration, they are guided by innate information about their direction of migration. It is generally assumed that this direction is represented twice, namely with respect to celestial rotation and with respect to the Earth's magnetic field^{1,2}. The interactions between the two cue systems have been analysed by exposing hand-raised young birds during the premigratory period to cue-conflict situations, in which celestial rotation and the magnetic field provided different information. Celestial rotation altered the course with respect to the magnetic field^{3–7}, whereas conflicting magnetic information did not seem to affect the course with respect to the stars^{8,9}. Celestial information thus seemed to dominate over magnetic information. Here we report that the interaction between the two cue systems is far more complex than this. Celestial rotation alone seems to provide only a tendency to move away from its centre (towards geographical south), which is then modified by information from the magnetic field to establish the distinctive, population-specific migratory direction.

Our experimental birds were garden warblers, *Sylvia borin*, a species of long-distance night migrants which breed in central and northern Europe and winter in tropical Africa. During migration, birds of the central European population first head southwest to Iberia, then change to a more southeasterly course to reach their African winter quarters¹⁰. Laboratory studies have indicated that the migratory direction of garden warblers is coded with respect to the magnetic field¹¹ and with respect to celestial rotation⁹. The change in direction from southwest to southeast was found to occur spontaneously at the end of September and could also be observed in the laboratory¹².

Our test birds were taken from their nests at the age of 4 to 6 days, transferred to Frankfurt am Main, and hand raised until they were self-sufficient. During the premigratory period, they were exposed to an artificial 'sky' with small lights as 'stars'^{9,13}, which rotated anti-clockwise with one rotation per day. Previous studies have shown that such a 'sky' could replace the natural sky in allowing stellar orientation⁹ and changing the magnetic course⁵. The cage containing the birds was positioned eccentrically so that the centre of rotation coincided with magnetic north, mimicking the natural situation. The control group experienced this 'sky' together with the local geomagnetic field (47 μ T, 66° inclination), whereas the experimental group experienced the sky in the absence of meaningful magnetic information, as the horizontal

component of the magnetic field was compensated (90° inclination). After these exposures, which lasted from mid-July to mid-August, the birds were moved into a windowless room and housed in individual cages.

Orientation tests were done between mid-August and mid-November under the same, but now stationary, sky in a vertical magnetic field (90° inclination). This means that during testing, neither celestial rotation nor magnetic information was available; the birds had to rely entirely on the stars and what they had learned during exposure. The mean directions of both groups during the first part and the second part of the season are shown in Fig. 1; Table 1 gives the data from individual birds.

During August and September, both groups showed significant orientation, but there was a difference between their headings. The controls that had been exposed to conditions simulating the natural situation showed the seasonally appropriate tendency to orient towards the southwest, whereas the experimental group that had observed the same rotating sky without magnetic information headed southwards. This difference in mean direction was significant ($P < 0.01$, Watson-Williams test¹⁴). From 1 October onwards, when garden warblers normally change their course towards the southeast, no such change in direction was observed. Both continued as before ($P > 0.05$, Wilcoxon test comparing the first half and the second half of the season) with a significant increase in scatter ($P < 0.05$, Mann-Whitney test).

Our findings indicate that the magnetic field is crucial for the development of the population-specific migratory direction with respect to the stars. Garden warblers must observe the rotating sky together with magnetic information during the premigratory period to establish, with respect to the stars, the specific south-

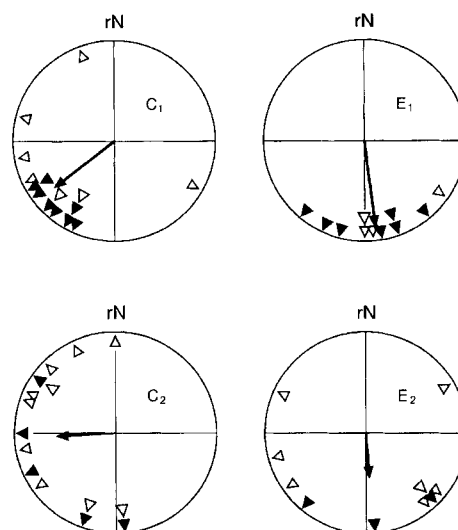


FIG. 1 Orientation of young garden warblers under a stationary artificial sky in the absence of meaningful magnetic information. Left, controls (C) that had been exposed, during the premigratory period, to the same, but rotating, sky in the geomagnetic field; right, experimental (E) birds that had been exposed to the same rotating sky in the absence of magnetic information. Top, August and September; bottom, October and November. The triangles indicate the mean directions of individual birds (filled, significant; open, not significant); the arrows represent the second-order mean vectors based on these directions proportional to the radius of the circle. rN, rotational north.

METHODS. The birds were tested in octagonal cages (1 m diameter, 35 cm high) with eight radially positioned double perches¹⁹. Their activity was recorded by switches connected to the perches. From the distribution of the activity, the heading of each individual test night was calculated by vector addition. A mean vector was calculated from the headings of each individual bird. The data were analysed using standard procedures of circular statistics¹⁴.

westerly course that they normally prefer during the first part of migration. This implies that the population-specific migratory direction, which deviates from due south, is not coded with respect to celestial rotation. Celestial rotation alone just seems to induce a tendency to move away from its centre, towards geographical south. Such a course would not be totally inappropriate under natural conditions, as it would make the birds head towards the equator to regions with less harsh conditions during the winter. However, it is very different from the population-specific migratory direction, which seems to be adapted to the geographical conditions met during the migration and helps species like the garden warbler to avoid ecological barriers such as the Alps, the Mediterranean Sea and the central part of the Sahara desert.

When garden warblers were raised by hand and later tested with the magnetic field as the only available cue, they preferred the seasonally appropriate southwesterly course and later changed to a south-southeasterly direction^{11,12}. This shows that the population-specific course for the first as well as for the second leg of the migration route is genetically coded with respect to the magnetic field. The directional behaviour of our control birds indicates that in the normal situation, with magnetic and celestial cues present, there are interactions between these two types of cues that set the population-specific course with respect to the stars. As a result, the birds are able to select that course with the help of stars alone.

Such a transfer of information from the magnetic field to celestial cues during the premigratory period, as suggested by our present data, seems to contradict previous findings. Former studies on the development of migratory orientation revealed

that, during this period, celestial rotation clearly dominates over information from the magnetic field. In cases of conflict, the innate magnetic compass course is modified and a new magnetic course is established which corresponds to the migratory direction with respect to celestial rotation³⁻⁷. Because, on the other hand, magnetic information is essential for setting the population-specific stellar course, it seems that celestial rotation does not simply override the magnetic course. The nature of the interactions between the two systems is thus not clear.

It is possible that the innate magnetic information does not specify a particular course, such as 230° southwest, but specifies a deviation from a reference point, something like '50° to the right' of a basic direction. This would mean that the migratory course is established in two stages. The birds would first locate a basic direction that may be provided by both cues, celestial rotation indicating geographical south, and the field lines of the geomagnetic field indicating magnetic south. In most cases, there will only be an insignificant difference between them. However, if there is a larger discrepancy, the birds would use the direction indicated by celestial rotation. In the second step, the innate magnetic information would define a certain deviation from this basic direction, resulting in the population-specific migratory course. The course thus established would then be remembered with respect to the magnetic field as well as with respect to the stars. Later, when the birds start to migrate, it could be located with either system.

The above model proposes that the basic direction indicating 'south' can be derived from celestial rotation as well as from the magnetic field, whereas the deviation from this basic direction defining the population-specific migration course is coded only

TABLE 1 Orientation of garden warblers by stars in the absence of magnetic information

Bird	August and September			October and November		
	<i>n</i>	<i>r</i>	α	<i>n</i>	<i>r</i>	α
Control group, exposed to a rotating 'sky' in the geomagnetic field						
1	11	0.80***	225°	14	0.67**	247°
2	10	0.49	341°	2	0.99	333°
3	11	0.90***	240°	7	0.86**	201°
4	13	0.49	122°	11	0.63*	306°
5	10	0.44	286°	10	0.54	176°
6	9	0.59	247°	12	0.52*	174°
7	8	0.68*	242°	13	0.22	1°
8	7	0.68*	235°	5	0.64	259°
9	7	0.65*	221°	5	0.88*	271°
10	7	0.20	262°	2	0.51	307°
11	7	0.88**	211°	3	0.79	292°
12	8	0.35	223°	5	0.50	200°
13	6	0.84*	212°	3	0.66	296°
14	7	0.44	205°	5	0.54	236°
15	6	0.90**	212°	5	0.68	315°
<i>N</i> = 15		0.64	232°, 0.76***		0.63	267°, 0.59**
Experimental group, exposed to a rotating 'sky' in the absence of magnetic information						
1	15	0.81***	142°	3	0.70	60°
2	12	0.68**	159°	13	0.72**	173°
3	11	0.77**	197°	15	0.88**	137°
4	11	0.44	174°	12	0.44	254°
5	13	0.75**	203°	2	0.66	294°
6	11	0.39	177°	13	0.46	237°
7	14	0.41	127°	10	0.45	141°
8	10	0.79**	170°	11	0.69**	222°
9	12	0.72**	158°	11	0.47	136°
10	10	0.79**	219°	0		
11	11	0.40	182°	11	0.62*	132°
<i>N</i> = 11 (10)		0.72	173°, 0.90***		0.64	176°, 0.46 ^{ns}

Orientation of individual garden warblers under stationary stars in the absence of meaningful magnetic information. *n*, tests per bird; α and *r*, direction and length of the bird's mean vector for the respective part of the season. Asterisks at *r* indicate significance by the Rayleigh test¹⁴. Significance levels: ns, not significant; *, *P* < 0.05; **, *P* < 0.01; ***, *P* < 0.001.

with respect to the magnetic field. There is an ecological advantage of preferably relying on celestial rotation for the basic direction. When innate information has to be transformed into an actual compass course, it is important that this results in a particular geographical direction. 'Away from the centre of rotation' always indicates true (geographical) south, whereas the local magnetic conditions may vary. Really strong magnetic anomalies are fairly rare; however, at higher latitudes, birds regularly meet large magnetic declinations (large deviations between magnetic and geographical north) and the most rapid changes in declination. Using celestial rotation as reference will thus help migrants to cope with unpredictable local magnetic conditions. It ensures that the magnetic course as well as the stellar course indicates the correct route when the birds start their migration. Indeed, experimental evidence so far indicates that celestial rotation dominates over information from the magnetic field when the migration course is first established at higher latitudes³⁻⁷. Later, during migration itself, when the birds have reached regions where magnetic declination is smaller and changes less rapidly, the magnetic field gains control¹⁵⁻¹⁷. We can only speculate why the population-specific deviation from due south is apparently encoded with respect to only the magnetic field. The direction derived from celestial rotation is a mental construct based on a rather complex integration of visual stimuli, whereas the magnetic field lines provide a stimulus that can be perceived directly. Possibly, it is easier to define angular deviations with respect to such a stimulus. The usefulness of the magnetic field for indicating angular differences is independent of the local relationship between magnetic and geographical north.

Another point emerges from our data: none of the groups showed the change in direction normally observed in the garden warbler at the end of September, when the free-flying birds leave Iberia for their African winter quarters. This is perhaps not surprising in the experimental group with their southerly tendencies, but the control group, after preferring a seasonally appropriate course during the first part of migration, might have been expected to change to a south-southeasterly heading. In previous studies with garden warblers, spontaneous changes in direction have been observed only when the birds were tested in the presence of suitable magnetic information^{12,18}, whereas birds tested under stars alone failed to show such a change⁹. Apparently, the population-specific course of the second part of migration also requires magnetic information. This suggests a transfer process analogous to that described for establishing the first direction. The behaviour of our control birds during October and November indicates that this process did not take place during the premigratory period when our test birds experienced celestial cues and the magnetic field together, but apparently occurs later, possibly shortly before the second part of migration begins. At that time, however, none of our test groups had access to magnetic information in the presence of stars. The increase in scatter observed in both groups might indicate a certain confusion caused by a lack of appropriate directional information. □

Received 9 April; accepted 18 July 1996.

1. Wiltschko, R. & Wiltschko, W. *Magnetic Orientation in Animals* (Springer, Berlin, 1995).
2. Able, K. P. & Able, M. A. *J. Exp. Biol.* **199**, 3-8 (1996).
3. Bingman, V. P. *Behaviour* **87**, 43-53 (1983).
4. Able, K. P. & Able, M. A. *Anim. Behav.* **39**, 905-913 (1990).
5. Able, K. P. & Able, M. A. *Nature* **347**, 378-379 (1990).
6. Able, K. P. & Able, M. A. *Nature* **364**, 523-525 (1993).
7. Prinz, K. & Wiltschko, W. *Anim. Behav.* **44**, 539-545 (1992).
8. Bingman, V. P. *Behaviour* **87**, 43-53 (1984).
9. Wiltschko, W., Daum, P., Fergenbauer-Kimmel, A. & Wiltschko, R. *Ethology* **74**, 285-292 (1987).
10. Zink, G. *Der Zug Europäischer Singvögel* Vol. 1 (Vogelzug, Möggingen, 1973).
11. Wiltschko, W. & Gwinner, E. *Naturwissenschaften* **61**, 406 (1974).
12. Gwinner, E. & Wiltschko, W. *J. Comp. Physiol.* **125**, 267-273 (1978).
13. Wiltschko, W. & Wiltschko, R. *J. Comp. Physiol.* **109**, 91-99 (1976).
14. Batschelet, E. *Circular Statistics in Biology* (Academic, New York, 1981).
15. Wiltschko, W. & Wiltschko, R. *Z. Tierpsychol.* **37**, 337-355 (1975).
16. Wiltschko, W. & Wiltschko, R. *Z. Tierpsychol.* **39**, 265-282 (1975).
17. Bingman, V. P. *Auk* **104**, 523-525 (1987).
18. Weindler, P. & Wiltschko, W. *Verh. Dtsch. Zool. Ges.* **84**, 369 (1991).
19. Wiltschko, W. *Z. Tierpsychol.* **25**, 537-558 (1968).

ACKNOWLEDGEMENTS. This work was supported by the Deutsche Forschungsgemeinschaft and by a scholarship of the Graduiertenförderung des Landes Hessen to P.W. We thank P. Berthold and the Vogelwarte Radolfzell for help in obtaining the test birds, and all who assisted in hand raising and keeping the birds.

CORRESPONDENCE and requests for materials should be addressed to W.W. (e-mail: wiltschko@zoology.uni-frankfurt.d400.de).

The dosage compensation system of *Drosophila* is co-opted by newly evolved X chromosomes

Ignacio Marín, Axel Franke, Greg J. Bashaw & Bruce S. Baker

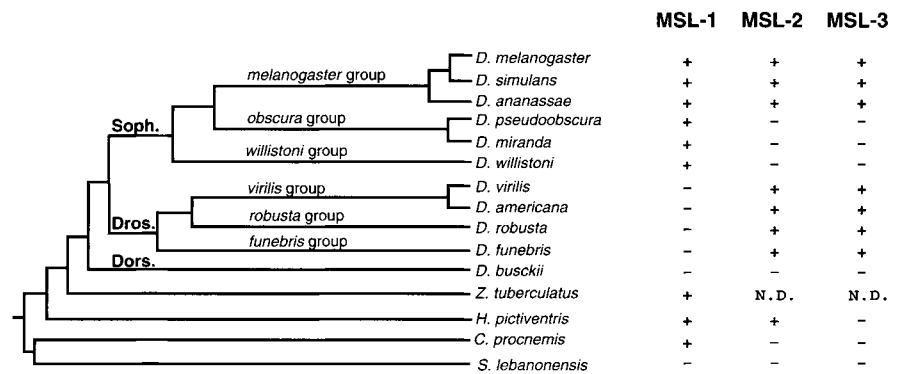
Department of Biological Sciences, Stanford University, Stanford, California 94305, USA

IN species where males and females differ in number of sex chromosomes, the expression of sex-linked genes is equalized by a process known as dosage compensation. In *Drosophila melanogaster*, dosage compensation is mediated by the binding of the products of the *male-specific lethal* (*msl*) genes to the single male X chromosome. Here we report that the sex- and chromosome-specific binding of three of the *msl* proteins (MSLs) occurs in other drosophilid species, spanning four genera. Moreover, we show that MSL binding correlates with the evolution of the sex chromosomes: in species that have acquired a second X chromosome arm because of an X-autosome translocation, we observe binding of the MSLs to the 'new' (previously autosomal) arm of the X chromosome, only when its homologue has degenerated. Moreover, in *Drosophila miranda*, a Y-autosome translocation has produced a new X chromosome (called neo-X), only some regions of which are dosage compensated. In this neo-X chromosome, the pattern of MSL binding correlates with the known pattern of dosage compensation.

Dosage compensation arises to deal with aneuploidy created by the degeneration of a sex-determining chromosome that occurs when recombination with its homologue is restricted^{1,2}. The evolutionary relationships of dosage compensation systems with the mechanisms that control sex are complex and poorly understood. Sex determination evolves rapidly; considering just dipterans, XX-X0, XX-XY, ZZ-ZW and more complicated systems involving several pairs of chromosomes have been described³. Two factors contribute to this rapid evolution: (1) regulatory changes: if a new gene takes control of the sex-determination pathway, the chromosome on which it is found becomes the sex-determining chromosome; and (2) chromosomal changes: for example, in XX-X0 organisms, a translocation between the X chromosome and an autosome leads to restriction of the homologue, non-translocated autosome, to males, as a new Y chromosome. The occurrence of either of these types of changes will eventually lead to the degeneration of the new sex-limited chromosomes, thereby generating a need for dosage compensation of their homologues. Thus, is the pre-existing dosage compensation machinery that functions on one chromosome recruited to fulfil the requirements created by a new chromosome degeneration? *A priori* there is no reason to expect such a recruitment, because dosage compensation is acquired on a gene-by-gene basis¹. Potentially, each gene could solve the problem in a different way. New mechanisms of dosage compensation could therefore arise to cope with the consequences of sex chromosome changes.

We have used a comparative approach to address this question

FIG. 1 Phylogenetic relationships among the species studied (based on refs 8, 11) and summary of results using the anti-MSL antibodies. The subgenera and groups are detailed for the *Drosophila* genus species (subgenera: Soph., *Sophophora*; Dros., *Drosophila*; Dors., *Dorsilopha*). Genera other than *Drosophila*: Z., *Zaprionus*; H., *Hirtodrosophila*; C., *Chymomyza*; S., *Scaptodrosophila*. +, male-specific staining in the arm or arms of the X chromosome; –, absence of sex-and chromosome-specific staining. MSL-1, MSL-2 and MSL-3 refer to the proteins against which the antibodies were raised (see Methods). N.D., not determined.



experimentally. Initially, we searched in other drosophilid species for the system first characterized in *Drosophila melanogaster*. In *D. melanogaster*, males are the heterogametic sex. Dosage compensation is mediated, for nearly all X-linked genes, by specific binding to the male X chromosome of the products of at least four genes, collectively named *male-specific lethals*: *msl-1*, *msl-2*, *msl-3* and *maleless* (*mle*). Binding of the MSL proteins increases the transcriptional activity of the genes on the male X chromosome to a level equal to that of the genes on both female X chromosomes⁴. Antibodies directed against the MSL-1, MSL-2 and MSL-3 proteins have been used for immunostaining in 14 drosophilid species. In 12 of these species, male X chromosome-specific staining by at least one of these antibodies is observed (Figs 1, 2). As in *D. melanogaster*, where female-specific repression of the production of MSL-2 protein prevents the function of the *msl*-based system^{5–7}, staining in females has not been observed in any of these 14 species. The male- and chromosome-specific binding of the MSL proteins provides strong evidence for the functional conservation of this system of dosage compensation in all of these species. Interestingly, staining is seen in both *Chymomyza* and

Hirtodrosophila, among the more distant relatives of *Drosophila* within the *Drosophilinae* subfamily⁸. The *msl*-based system of dosage compensation is therefore much older than the *Drosophila* genus; the *Chymomyza*–*Drosophila* split is estimated to have occurred at least 55 million years ago^{9–11}.

Using the anti-MSL antibodies, we have been able to examine whether the MSL-based system is responsible for dosage compensation in situations where structural changes in the sex chromosomes have occurred. Occasionally during the diversification of the *Drosophila* genus, X-autosome translocations have been fixed, generating new X chromosome arms (Fig. 3). In *D. americana*, the homologue of the translocated arm has not yet degenerated, therefore it is not expected that dosage compensation will be required in the new X chromosome arm. In this species, MSL binding is observed in a single chromosomal arm (Fig. 4a, b). In three independent cases, where the homologue of the translocated (ex-autosomal) arm is fully degenerated, binding of the MSL proteins to the new arm of the X chromosome is observed (*D. pseudoobscura*, *D. willistoni* and *D. robusta*; see Fig. 4c–h). The number of sites stained is similar in both arms and is comparable

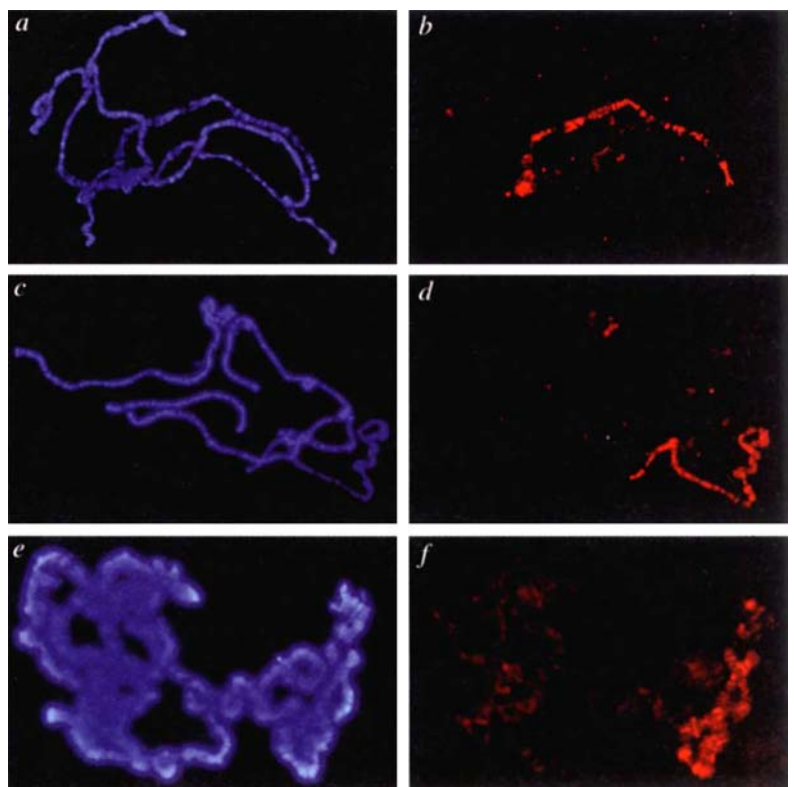


FIG. 2 Immunostaining in males of several drosophilid species. Left panels, DAPI staining; right panels, anti-MSL staining. a, b, *D. virilis* (the right panel shows the anti-MSL-3 antibody signal). c, d, *Zaprionus tuberculatus* (anti-MSL-1). e, f, *Chymomyza procnemis* (anti-MSL-1).

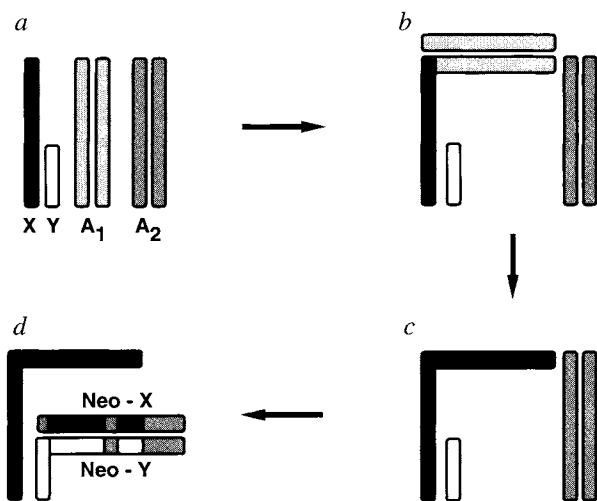


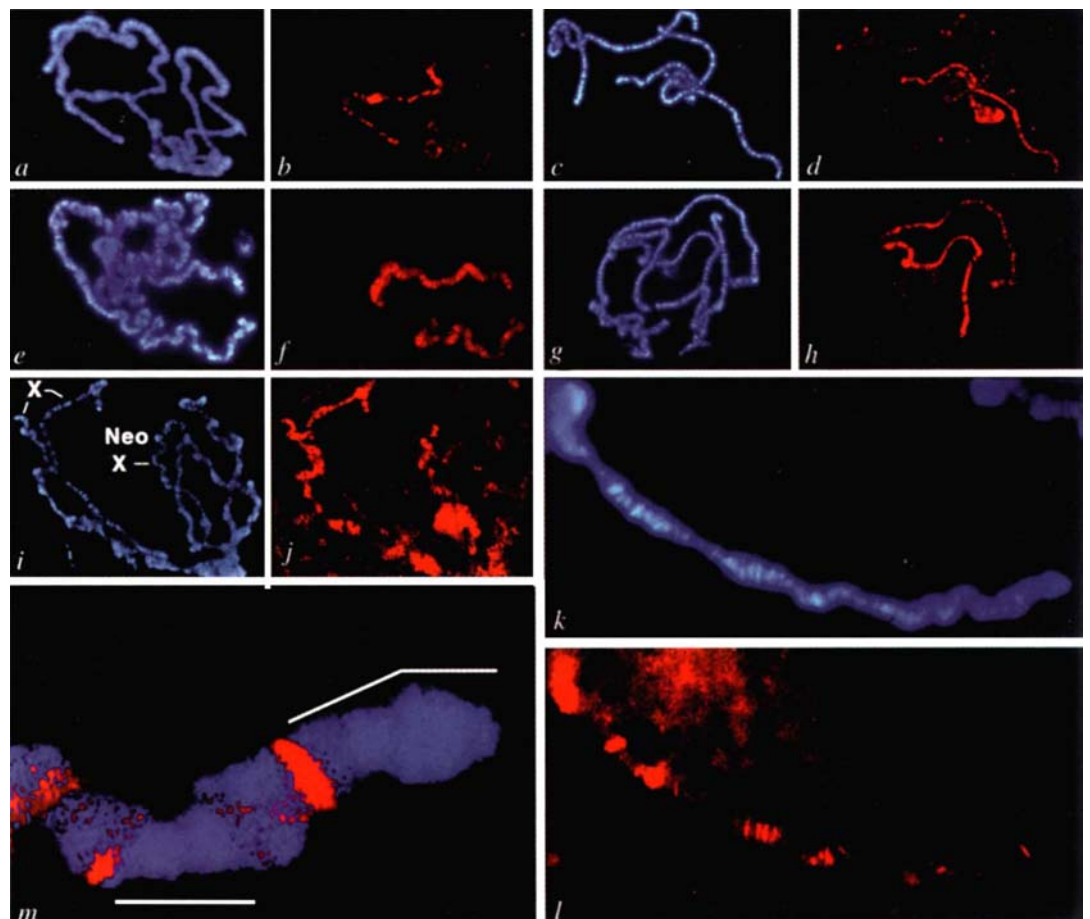
FIG. 3 Sex chromosome evolution in selected species of the *Drosophila* genus. Arrows refer to the sequential steps in chromosomal evolution; no phylogenetic relationship among species is implied. *a*, Simplified male karyotype of most *Drosophila* species. Only three of the five pairs of chromosomes are shown. The X chromosome is dosage compensated (black), whereas the Y chromosome is inactive (white). *A*₁ and *A*₂ are two pairs of autosomes. *b*, Karyotype found when an X-autosome translocation is fixed and the non-translocated arm has not yet degenerated (for example, *D. americana*). *c*, Karyotype after the homologue of the translocated chromosome has degenerated (for example, *D. willistoni*, *D. pseudoobscura* and *D. robusta*). *d*, In *D. miranda*, a recent (at most two million years old⁹) Y-autosome translocation has occurred, in addition to the X-autosome translocation characteristic of *D. pseudoobscura* and its relatives¹⁶. The autosomal arm that has been fused to the Y chromosome (neo-Y) is partially degenerated (white patches) and its homologue (neo-X) is partially dosage compensated (shown as black patches)^{13,14}.

to that observed in *D. melanogaster*. Previous experiments demonstrated that in *D. willistoni* and *D. pseudoobscura*, both arms of the X chromosome (shown in black) were fully compensated¹². In *D. miranda*, a recent Y-autosome translocation has occurred (Fig. 3), in addition to the X-autosome translocation also observed in its sibling *D. pseudoobscura*. The neo-Y chromosome is not fully degenerated, and some regions of its homologue (neo-X, also known as X₂) do not show increased transcriptional activity^{13,14}. In the neo-X chromosome there are fewer sites showing MSL binding than in the arms of the X chromosome, and the zones with binding correlate with those zones known to be dosage

compensated (Fig. 4*i–m*). These results are expected if, as suggested for *D. melanogaster*⁴, the *msl*s are the direct mediators of dosage compensation in all these species.

Our results have several implications. First, they show that the *msl*-based system is older than the *Drosophila* genus. Second, because this system can be co-opted to new chromosomes during evolution, it may be older than the degeneration process that causes a requirement for dosage compensation in new chromosomal arms. Therefore, the *msl*-based system may operate in diverse organisms, even if their sex determination systems are different. Third, in four independent cases, the solution to the

FIG. 4 Immunostainings showing the correlation between degeneration of a chromosome and acquisition of MSL binding in its homologue. *a, c, e, g, i, k*, DAPI staining; *b, d, f, h, j, l*, antibody staining. *a, b*, *D. americana* (anti-MSL-3 antibody); binding is present only in one of the arms of the X chromosome. *c, d*, *D. pseudoobscura* (anti-MSL-1). *e, f*, *D. willistoni* (anti-MSL-1). *g, h*, *D. robusta* (anti-MSL-3). In *c–h*, where the non-translocated homologue has degenerated, both arms show binding. *i–m*, *D. miranda* (anti-MSL-1 antibody). *i, j*, General view. Three chromosomal arms, both arms of the X plus the neo-X, are stained. *k, l*, Close-up of the neo-X chromosome of *D. miranda*. The number of bands, especially near the telomeric end, is not as abundant as usually observed in the X chromosomes. *m*, Anti-MSL-1 staining (red) superimposed on the DAPI staining (blue) in the tip of the neo-X chromosome of *D. miranda*. The white lines correspond to the zones determined to be non-dosage compensated¹⁴.



problem of sex chromosome degeneration has been to extend the *msl*-based mechanism. This has occurred despite the fact that, on an evolutionary time scale, the requirements for compensation appear gradually, gene by gene. These results suggest that the extant compensatory system is adopted when new requirements for dosage compensation arise. The constraints on developing new dosage compensation mechanisms probably derive from the difficulty in evolving systems able sex-specifically to regulate functionally unrelated genes, which have in common only the fact that they reside on the same chromosome. Taking all these considerations together, we expect to find identical mechanisms functioning in phylogenetically very distant species.

Note added in proof: Histone H4 acetylated at lysine residue 16, known to be implicated in the hypertranscription mechanism in *D. melanogaster*, is also enriched in the newly evolved X chromosomes of *D. pseudoobscura* and *D. miranda*¹⁷. □

Methods

Antibodies. We have used rat polyclonal antibodies against three of the MSLs (MSL-1, amino acids 618–939; MSL-2, 534–669 (ref. 5); and MSL-3, 324–513 (ref. 15)). These antibodies were selected on the basis of results of preliminary experiments using *Drosophila pseudoobscura*, *D. virilis* and *Chymomyza procnebris*. Only these antibodies stained the male X chromosomes of at least one of these species. Six other antibodies, including two raised against MLE, gave negative results. The results summarized in Fig. 1 are based on at least two preparations for each combination of antibody, sex and species. The negative assessments for *Scaptodrosophila lebanonensis* and *Drosophila busckii* males are based on at least 5 chromosome preparations per antibody. The number of epitopes recognized in distant species is probably small because: (1) none of the antibodies is positive in all the species; (2) positive signals do not strictly correlate with the phylogenetic proximity to *D. melanogaster*; and (3) signal intensity in other species is lower than in *D. melanogaster*. Therefore, the simplest explanation for the failure to observe binding in those two species is that they have lost the conserved epitope(s) recognized by our antibodies.

Immunostainings. Third instar larvae were sexed according to the size of their gonads. Sex was verified by observation of polytene chromosomes from the salivary glands. Under phase contrast optics, males showed one to three pale arms, depending on the species. Moreover, in some cases the male X chromosome had a peculiar shape, was shorter or seemed diffuse when compared with the autosomes. When possible, it was further determined that the stained chromosome(s) corresponded to the X chromosome(s) by comparison with the available chromosome maps. Immunostainings were performed as described in ref. 5. Antibodies were generally used at the same dilution routinely chosen for *D. melanogaster* (1:50). Although at this concentration the intensity of staining observed in the other species was substantially lower than in *D. melanogaster*, the male X chromosome was easily visualized in every preparation of most species. Only in *D. pseudoobscura*, *D. miranda* and *Hirtodrosophila pictiventris* was the fluorescence level so low that in some slides the X chromosome was not easily distinguishable. In these species, higher concentrations (up to 1:20) were used. DAPI was used to stain the chromosomes for localization and study under fluorescence optics.

Received 1 May; accepted 7 August 1996.

- Charlesworth, B. *Science* **251**, 1030–1033 (1991).
- Charlesworth, B. *Curr. Biol.* **6**, 149–162 (1996).
- Bull, J. J. *Evolution of Sex Determining Mechanisms* (Benjamin/Cummings, 1983).
- Baker, B. S., Gorman, M. & Marin, I. *Annu. Rev. Genet.* **28**, 491–521 (1994).
- Bashaw, G. J. & Baker, B. S. *Development* **121**, 3245–3258 (1995).
- Zhou, S. et al. *EMBO J.* **14**, 2884–2895 (1996).
- Kelley, R. L. et al. *Cell* **81**, 867–877 (1995).
- Grimaldi, D. A. *Bull. Am. Mus. Nat. Hist.* **197**, 1–139 (1990).
- Russo, C. A. M., Takezaki, N. & Nei, M. *Mol. Biol. Evol.* **12**, 391–404 (1995).
- Beverley, S. M. & Wilson, A. C. *J. Mol. Evol.* **21**, 1–13 (1984).
- Kwiatkowski, J., Skarecky, D., Bailey, K. & Ayala, F. J. *J. Mol. Evol.* **38**, 443–454 (1994).
- Abraham, I. & Lucchesi, J. C. *Genetics* **78**, 119–126 (1974).
- Strobel, E., Pelling, C. & Arnheim, N. *Proc. Natl Acad. Sci. USA* **75**, 931–935 (1978).
- Das, M., Mutsuddi, D., Duttagupta, A. K. & Mukherjee, A. S. *Chromosoma* **87**, 373–388 (1982).
- Gorman, M., Franke, A. & Baker, B. S. *Development* **121**, 463–475 (1995).
- Dobzhansky, T. *Genetics* **20**, 377–391 (1935).
- Steinemann, M., Steinemann, S. & Turner, B. M. *Chromosome Res.* **4**, 185–190 (1996).

ACKNOWLEDGEMENTS. We thank D. A. Grimaldi for his advice concerning drosophilid phylogeny; G. Bohm for preparation of fly food; and the National *Drosophila* Species Resource Center (Bowling Green, Ohio) and the European *Drosophila* Stock Centre (Umea, Sweden) for *Drosophila* stocks. This work was supported by a postdoctoral fellowship (Ministerio de Educación y Ciencia, Spain) to I.M., an NIH predoctoral training grant to G.J.B. and an NIH grant to B.S.B.

CORRESPONDENCE and requests for materials should be addressed to I.M. (e-mail: marin@cmgm.stanford.edu).

Functional neuroanatomy of human rapid-eye-movement sleep and dreaming

Pierre Maquet*†, Jean-Marie Péters*, Joël Aerts*, Guy Delfiore*, Christian Degueldre*, André Luxen* & Georges Franck*†

* Cyclotron Research Centre (B30), University of Liège, 4000 Liège, Belgium

† Department of Neurology, CHU Sart Tilman (B35), 4000 Liège, Belgium

RAPID-EYE-MOVEMENT (REM) sleep is associated with intense neuronal activity, ocular saccades, muscular atonia and dreaming^{1,2}. The function of REM sleep remains elusive and its neural correlates have not been characterized precisely in man. Here we use positron emission tomography and statistical parametric mapping to study the brain state associated with REM sleep in humans. We report a group study of seven subjects who maintained steady REM sleep during brain scanning and recalled dreams upon awakening. The results show that regional cerebral blood flow is positively correlated with REM sleep in pontine tegmentum, left thalamus, both amygdaloid complexes, anterior cingulate cortex and right parietal operculum. Negative correlations between regional cerebral blood flow and REM sleep are observed bilaterally, in a vast area of dorsolateral prefrontal cortex, in parietal cortex (supramarginal gyrus) as well as in posterior cingulate cortex and precuneus. Given the role of the amygdaloid complexes in the acquisition of emotionally influenced memories, the pattern of activation in the amygdala and the cortical areas provides a biological basis for the processing of some types of memory during REM sleep.

Thirty young, healthy, right-handed male subjects (mean age, 22.5 years; range, 20–25) participated in this study, which took place over 3 polygraphically monitored nights spent on a scanner couch, at one-week intervals. Polygraphic recordings included electroencephalogram (recorded between electrode pairs C3–A2 and C4–A1), EOG (electro-oculogram) and chin EMG (electromyogram), and were scored using the criteria of ref. 3. Subjects were asked not to sleep during the night preceding the second and third test nights to ensure optimal sleep stability despite difficult experimental conditions (head stabilized using a face mask secured to the scanner head holder and left arm cannulated and immobilized). Subjects were selected if they could maintain two periods of slow-wave sleep and of REM sleep (20 min each) during the first two test nights. Nineteen subjects were finally selected.

During the third test night, regional cerebral blood flow (rCBF) distribution was recorded as an index of neuronal activity using six slow intravenous infusions of H₂¹⁵O scheduled during the following states: wakefulness, slow-wave sleep, slow-wave sleep, REM sleep, REM sleep, wakefulness. This order was imposed by the physiological preponderance of slow-wave sleep early in the night and of REM sleep in the early morning. During sleep, scans were done when polysomnography showed characteristic sleep patterns (slow-wave or REM sleep) that remained stable over the 5-min period necessary for water production. After each sleep scan, the subjects were woken up and asked to describe what was in their minds. Dream reports were scored following the classification system of ref. 4. Four subjects were studied during a complete set of states (2 wakeful, 2 slow-wave, 2 REM sleep states); another group of three subjects was scanned during 2 wakeful and 2 REM sleep states but changed their sleep stage during the injections in the slow-wave sleep state. All subjects recalled a dream after awakening from REM sleep. The dream contents included various settings, characters, objects and activities. A final group of four subjects were scanned during 2 wakeful and 2 slow-wave sleep

states but changed their sleep stage during the injections in the REM sleep stage. The analysis was done using these three groups (11 subjects total, 7 with stable REM sleep). Positron emission tomography (PET) data were analysed using statistical parametric mapping. Correlations between rCBF and the presence of REM sleep were assessed (see Fig. 1 legend), to identify brain areas whose activity was associated with REM sleep.

Significant positive correlations between rCBF and REM sleep (Table 1 and Fig. 1) were observed in the anterior cingulate cortex (mainly Brodmann's area (BA) 24), the posterior part of the right parietal operculum (anterior part of BA 40), the right amygdala and surrounding entorhinal cortex, and in a region encompassing the left amygdala and surrounding entorhinal cortex, the thalamic nuclei, the dorsal mesencephalon and pontine tegmentum. Within this area, local maxima of blood flow were located in the left amygdala, left thalamus and pontine tegmentum. Significant negative correlations between rCBF and REM sleep (Table 2 and Fig. 2) were found in precuneus and adjacent posterior cingulate cortex (BA 31) and, bilaterally, in a large area of dorsolateral prefrontal cortex (mainly BA 10, but also 46, 9, 8 and the lateral part of BA 11) and in parietal cortex (supramarginal gyrus, BA 40).

These results provide the first comprehensive description of the distribution of cerebral activity during REM sleep in man. They confirm previous PET data suggesting a metabolic activation of the brainstem⁵, the limbic system⁶ and the cingulate cortex⁷ during REM sleep. Recently, a positive correlation was reported between cerebral glucose metabolism and ocular saccades during REM sleep in regions where we report a negative correlation between rCBF and REM sleep (prefrontal and parietal areas)⁸. The results do not necessarily conflict with ours because the different variables characterized two different neural networks.

REM sleep is generated by nuclei of the mesopontine reticular formation¹⁹. These nuclei are an important source of activation of thalamic nuclei which, in turn, activate the cortical mantle⁹. The significant correlation between rCBF and REM sleep observed in pontine tegmentum and thalamic nuclei most probably reflects the activation of these central-core structures during REM sleep. The left-sided predominance of the thalamic activation may be due to a statistical bias of the sampled population, or may reflect true lateralization of thalamic activation in right-handed subjects.

Our results show that the intense and widespread cortical activation during REM sleep is not uniform: several regions are more, and others are less, active than the rest of the brain. The significant correlation found in the amygdala is conspicuous. The role of the amygdaloid complexes in REM sleep phenomenology, although suggested in animals¹¹, has never been explicitly demonstrated in humans. The amygdaloid complexes receive an important set of afferent nerve connections arising from the dorsal pons¹² and intralaminar thalamic nuclei¹³, by which they could be activated during REM sleep. On the other hand, the amygdaloid complexes project back to subcortical structures (hypothalamus and several brainstem nuclei)¹². By these subcortical projections, the amygdaloid complexes could induce the autonomic manifestations that are commonly observed during REM sleep¹⁵.

The amygdaloid complexes also have reciprocal connections with many cortical areas^{12,16}, which enable the amygdala to establish cross-modal sensory-affective associations^{12,16,17}. The amygdaloid complexes are also implicated in the formation and consolidation of memories paired with emotional stimuli^{12,16}, by regulating the storage of long-term memory in other cerebral regions¹⁸. Periods of REM sleep have also been implicated in the consolidation of memory traces¹⁹. In man, REM sleep particularly influences the processing of 'procedural-implicit' memory¹⁹ and also of memories acquired under emotionally charged conditions²⁰.

TABLE 1 Increased brain activity during REM sleep

Area (Brodmann's area)	x	y	z	Z score	P (corrected)
Anterior cingulate cortex (24)	2	8	36	5.15	0.001
	0	-12	44	4.47	0.010
Right parietal operculum (40)	60	-22	16	4.88	0.002
Right amygdaloid complex	16	-4	-12	4.29	0.021
Left amygdaloid complex	-18	-6	-12	5.15	0.001
Left thalamus	-10	-22	4	4.05	0.049
Pons	-6	-34	-28	4.21	0.028

Localization and statistical results concerning the local maxima of the brain areas where the activity is positively correlated with REM sleep. Coordinates are defined in the stereotactic space of Talairach²⁸, relative to anterior commissure. x represents the lateral distance from midline (positive, right); y is the anteroposterior distance from anterior commissure (positive: anterior); z represents the rostrocaudal distance from the bicommissural plane (positive: rostral). The areas are significantly correlated at a threshold of $P = 0.001$, by reference to unit normal distribution ($Z = 3.09$), and at a threshold of corrected $P < 0.05$ (corrected P , the probability that the rCBF variation could have occurred by chance over the entire volume analysed).

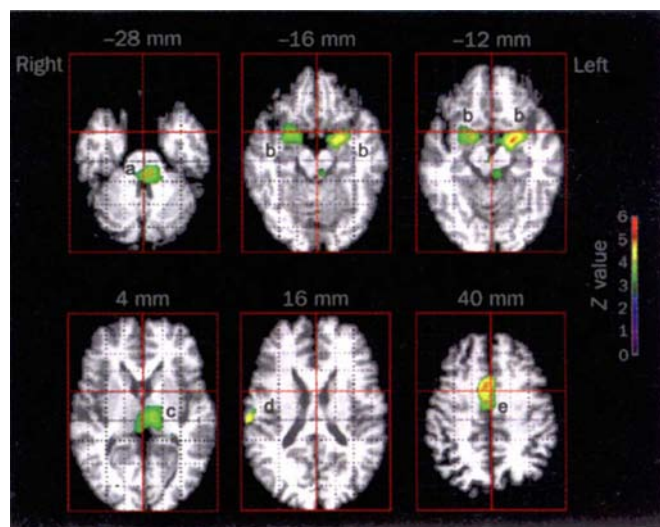


FIG. 1 Transverse sections showing brain areas where activity showed a significant positive correlation with REM sleep. Functional PET results are displayed at threshold of $Z = 3.09$ ($P < 0.001$), and superimposed, for anatomical reference, upon a T1-weighted magnetic resonance imaging scan normalized into the Talairach space²⁸. Section numbers refer to the distance from bicommissural plane. a, pons; b, amygdaloid complexes; c, thalamus; d, right parietal operculum; e, anterior cingulate cortex.

METHODS. PET acquisitions were obtained with a Siemens CTI 951 scanner in 2D mode. For each subject, six 120-s scans were made using $H_2^{15}O$ infusion technique, under continuous polygraphic monitoring. PET data were corrected for attenuation and background activity, and reconstructed with a Hanning filter (cutoff frequency, 0.5 cycles per pixel; 8.7 mm FWHM (full width at half-maximum)). Data were processed using the statistical parametric mapping software. Individual data were realigned²⁸, normalized into stereotactic space²⁸ then smoothed with a 12-mm FWHM gaussian filter. A design matrix was specified, which included global activity as confounding covariate²⁹. This matrix allowed for an across-group analysis looking for regionally specific condition effects. The analysis identified the brain regions where rCBF was significantly correlated with the presence of a particular functional state (such as REM sleep), using linear contrasts (+1 or -1 for REM sleep scans, 0 for slow-wave sleep and wakeful-state scans). The resulting maps (SPM{t}) were transformed to the unit normal distribution (SPM{Z}) and thresholded at $P < 0.001$ ($Z = 3.09$). The resulting foci of activation were characterized in terms of the probability that the peak height could have occurred by chance over the entire volume analysed (corrected P value ≤ 0.05). Details of the wakeful and slow-wave sleep states will be published elsewhere.

Thus, the activation of the amygdaloid complexes suggests that REM sleep contributes to some memory processing. Moreover, the rCBF distribution suggests a functional link between the amygdala, the hippocampal formations and cortical areas during REM sleep. First, the anterior part of the entorhinal cortex is embedded within the activation foci centred on the amygdaloid complexes. As they have dense reciprocal connections with the entorhinal cortex¹², their coactivation is possible. Second, the anterior cingulate cortex to which the amygdaloid complexes send their largest cortical efferent projection¹², is also significantly correlated with REM sleep. Third, a significant correlation is found in the right parietal operculum. Although this activation might be due to a statistical bias of the sampled population, we feel that it corresponds to an actual asymmetrical activation. In non-human primates, this area is the only somatosensory region that receives amygdalar projections¹⁶. Its activation might reflect the processing of the somesthetic stimulations, common to all our subjects, coming from the cannulation and immobilization of their left arm.

Other cortical areas receive abundant amygdalar projections (medial orbitofrontal, temporal, occipital and insular cortices)^{12,16}. The activation of occipital^{8,21} and temporal⁵ areas during REM sleep have been suggested by previous neuroimaging studies in humans. Here, none of these regions were significantly correlated with REM sleep, probably because the group analysis averaged out cortical activations specific to each individual subject or even to each scanned REM sleep episode.

A further indication of amygdalocortical interplay during REM sleep is, that the regions which are negatively correlated with REM sleep are known to receive few or no amygdalar efferents^{12,16,22,23}, and consequently, are not likely to be activated during REM sleep as much as cortical areas characterized by high amygdalar input. Dorsolateral prefrontal areas and precuneus have been shown to participate in the encoding and retrieval of episodic memory²⁴, whereas inferior parietal lobules (and pre-

TABLE 2 Decreases in brain activity during REM sleep

Area (Brodmann's area)	x	y	z	Z score	P (corrected)	
Right dorsolateral prefrontal cortex (10)	22	56	-4	6.36	<0.001	
	36	48	0	5.86	<0.001	
	26	48	16	5.07	0.001	
	26	40	20	4.81	0.003	
	(11)	18	42	-12	4.34	0.017
	22	30	-16	4.19	0.029	
	(9)	20	38	36	4.78	0.003
	(46)	40	16	24	4.22	0.026
	(8)	28	10	48	5.32	<0.001
Left dorsolateral prefrontal cortex (10)	-20	44	-8	5.10	0.001	
	-34	48	-4	4.80	0.003	
	-26	54	4	4.62	0.006	
	-10	62	16	4.53	0.008	
	-16	56	-4	4.47	0.010	
	-16	62	8	4.44	0.012	
	(11)	-12	36	-16	4.94	0.001
	(46)	-32	44	16	5.27	<0.001
	(47)	-34	40	-8	5.16	<0.001
Right parietal cortex	(40)	48	-46	36	5.51	<0.001
		40	-68	36	4.45	0.011
Left parietal cortex	(40)	-44	-50	36	4.07	0.047
Precuneus and		8	-64	28	4.80	0.003
posterior cingulate cortex	(31)	16	-66	32	4.73	0.004
		2	-42	32	4.48	0.010

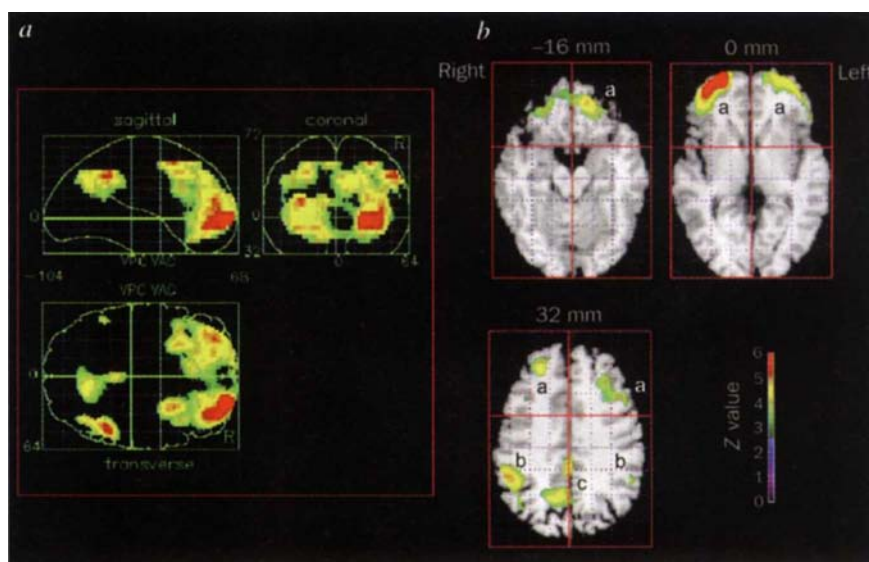
Localization and statistical results concerning the local maxima of the brain areas where the activity is negatively correlated with REM sleep. Coordinates and correlations are determined as for Table 1.

frontal areas) participate in working memory^{25,26}. Our results suggest that these memory processes are not prominent during REM sleep.

In summary, there are indications that during REM sleep, functional interactions occur between the amygdaloid complexes, anterior cingulate cortex and various, mainly posterior, areas. These interactions might lead to the reactivation of affective components of memories, thereby bringing about the consolidation of memory traces.

As all our subjects recalled a dream after REM sleep scans, the distribution of rCBF may reflect their dreaming activity. The amygdalo^{12,16} cingulate²⁷ coactivation could account for the emotional and affective aspects of dreams. Other formal char-

FIG. 2 Brain areas where activity showed a significant negative correlation with REM sleep. Functional PET results are displayed at threshold of $Z = 3.09$ ($P < 0.001$) by reference to the unit normal distribution. *a*, Projections within the stereotactic space²⁸. *b*, Transverse sections where functional data are superimposed, upon a T1-weighted MRI scan normalized into stereotactic space. Section numbers refer to the distance from bicommissural plane. Methods as in Fig. 1: *a*, prefrontal cortex; *b*, parietal areas; *c*, precuneus and adjacent posterior cingulate cortex.



acteristics of dreams (temporal distortions, weakening of self-reflective control, amnesia on awakening²) might be related to the relative prefrontal deactivation. The suggested amygdalo-cortical activation could account for the perceptual components of dreaming. □

Received 17 April; accepted 8 July 1996.

1. Jones, B. E. *Neuroscience* **40**, 637–656 (1991).
2. Hobson, J. A. *The Dreaming Brain* (Basic books, New York, 1988).
3. Rechtschaffen, A. & Kales, A. A. *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects* (Brain Information Service/Brain Research Institute, Los Angeles, 1968).
4. Hall, C. S. & van de Castle, R. L. *The Content Analysis of Dreams* (Appleton Century Crofts, New York, 1966).
5. Maquet, P. et al. *Brain Res.* **513**, 136–143 (1990).
6. Braun, A. R. et al. *Neurology* **42** (Suppl. 3), 182 (1992).
7. Buchsbaum, M. S. et al. *Life Sci.* **45**, 1349–1356 (1989).
8. Hong, C. C. H., Gillin, J. C., Dow, B. M., Wu, J. & Buchsbaum, M. S. *Sleep* **18**, 570–580 (1995).
9. Steriade, M., Jones, E. G. & Llinas, R. *Thalamic Oscillations and Signalling* (Wiley, New York, 1990).
10. Llinas, R. & Paré, D. *Neuroscience* **44**, 521–535 (1991).
11. Calvo, J. M., Badillo, S., Morales-Ramirez, M. & Palacios-Salas, P. *Brain Res.* **403**, 22–30 (1987).
12. Amaral, D. G., Price, J. L., Pitkanen, A. & Carmichael, S. T. in *The Amygdala. Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction* (ed. Aggleton, J. P.) 1–66 (Wiley-Liss, New York, 1992).

13. Aggleton, J. P., Burton, M. J. & Passingham, R. E. *Brain Res.* **190**, 347–368 (1980).
14. Davis, M. *Annu. Rev. Neurosci.* **15**, 353–375 (1992).
15. Orem, J. & Keeling, J. in *Physiology in Sleep* (eds Orem, J. & Barnes, C. D.) 315–335 (Academic, London, 1980).
16. Amaral, D. G. & Price, J. L. *J. Comp. Neurol.* **230**, 465–496 (1984).
17. Bechara, A. et al. *Science* **269**, 1115–1118 (1995).
18. Gallagher, M. & Chiba, A. A. *Curr. Opin. Neurobiol.* **6**, 221–227 (1996).
19. Smith, C. *Behav. Brain Res.* **69**, 137–145 (1995).
20. Smith, C. & Lapp, L. *Sleep* **14**, 325–330 (1991).
21. Madsen, P. L. et al. *J. Cereb. Blood Flow Metab.* **11**, 502–507 (1991).
22. Aggleton, J. P. *Trends Neurosci.* **16**, 328–333 (1993).
23. Olson, C. R. & Musil, S. Y. *J. Comp. Neurol.* **324**, 237–260 (1992).
24. Fletcher, P. C. et al. *Brain* **118**, 401–416 (1995).
25. Paulesu, E., Frith, C. D. & Frackowiak, R. S. J. *Nature* **362**, 342–345 (1993).
26. Jonides, J. et al. *Nature* **363**, 623–625 (1993).
27. Devinsky, O. et al. *Brain* **118**, 279–306 (1995).
28. Friston, K. J. et al. *Human Brain Mapp.* **2**, 165–189 (1995).
29. Friston, K. J. et al. *Human Brain Mapp.* **2**, 189–210 (1995).

ACKNOWLEDGEMENTS. We thank R. S. J. Frackowiak and K. J. Friston for providing the statistical parametric mapping software, P. Hawotte and M. Pierrard for (nocturnal) technical assistance, G. Hartstein for checking the manuscript, and S. Fuchs and V. Delvaux for help with subject selection. P.M. is research associate at the Fonds National de la Recherche Scientifique de Belgique (FNRS). This research was supported by FNRS grants, by the European Sleep Research Society-Synthelabo Research Grant 1994 and the Queen Elisabeth Medical Foundation.

CORRESPONDENCE and requests for materials should be addressed to P.M. (e-mail: maquet@pet.crc.ulg.ac.be).

Induction of cell death by endogenous nerve growth factor through its p75 receptor

José María Frade, Alfredo Rodríguez-Tébar* & Yves-Alain Barde

Max-Planck Institute for Psychiatry, Department of Neurobiochemistry, 82152 Planegg-Martinsried, Germany

* Cajal Institute, CSIC, Doctor Arce 37, E-28002 Madrid, Spain

DURING development, neuronal survival is regulated by the limited availability of neurotrophins, which are proteins of the nerve growth factor (NGF) family. Activation of specific *trk* tyrosine kinase receptors by the neurotrophins blocks programmed cell death. The *trkA*-specific ligand NGF has also been shown to activate the non-tyrosine kinase receptor p75, a member of the tumour necrosis factor (TNF) receptor and Fas (APO-1/CD95) family. Here we report that, early in development, endogenous NGF causes the death of retinal neurons that express p75 but not *trkA*. These results indicate that, as with cells of the immune system, the death of neurons in the central nervous system can also be induced by ligands, and that the effect of NGF on cell fate depends on the type of receptor expressed by developing neurons.

The elimination of neurons has long been recognized as an integral part of normal development. In vertebrates, this phenomenon is particularly well documented during the time corresponding to the innervation of target cells¹, and is easily monitored as net cell numbers decrease over time. During this period, programmed cell death can be prevented by the administration of exogenous neurotrophins, or increased by neutralizing the neurotrophins with antibodies or by deleting their genes². Prevention of cell death involves tyrosine kinase receptors of the *trk* family, as the deletion of the genes coding for these receptors generates phenotypes similar to those resulting from ligand elimination³. In addition, cell death can also be observed in the nervous system before innervation of the target cells. Soon after leaving the cell cycle, many neurons die, as has been shown in the cerebral cortex⁴, and in the worm *Caenorhabditis elegans*⁵. Here we consider the regulation of this early form of neuronal death.

In the absence of *trkA*, NGF can signal in non-transformed cells

through the neurotrophin receptor p75, as shown by the nuclear translocation of the transcription activator NF- κ B⁶. To examine the relevance of p75 activation by NGF *in vivo*, we investigated the consequences of NGF removal on the development of the chick retina. Previous *in situ* hybridization studies⁷ have demonstrated the presence of p75 messenger RNA as early as embryonic day 4 (E4), and staining of E6 retinal sections with antibodies to chick p75 confirmed the presence of this receptor mostly in the central retina, but also on scattered cells (data not shown). The *NGF* gene is expressed as early as E4, but the *trkA* gene is expressed only later in development (Fig. 1). Cell death is widespread in the dorsal part of the chick retina as early as E4 (ref. 8), before innervation of the tectum begins, and TUNEL staining, which detects DNA fragments on tissue sections, revealed a pattern corresponding to that of p75 expression (Fig. 2e). Double labelling with p75 antibodies showed that 95% of the TUNEL-positive cells were also p75-positive (94 cells counted). To test the possibility that endogenous NGF causes cell death, cells secreting an anti-NGF monoclonal antibody⁹ were applied onto chick embryos at E2.5. When examined at E6, the retinae of embryos treated with anti-NGF antibodies showed a remarkable reduction in the number of TUNEL-positive cells when compared with those of control embryos (Fig. 2; compare *c* with *d*, and *e* with *f*). To quantify the reduction in cell death resulting from the administration of anti-NGF antibodies, soluble nucleosomes were measured in retinal extracts by using an immunoassay combining anti-DNA and anti-histone antibodies¹⁰. A marked reduction was observed at E6 (Table 1), and in three embryos, cell death was reduced by as much as 80%. Similar effects were observed with sheep anti-NGF polyclonal antibodies, but hybridoma cells producing an anti-NT-3 antibody had no

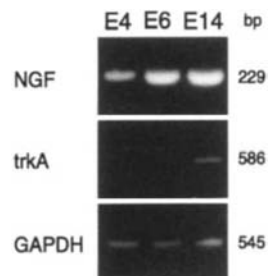


FIG. 1 NGF and *trkA* gene expression in the developing chick retina.

effect on developmental cell death (data not shown). NT-3 is also expressed in the early chick retina (like NGF), and is needed during retinogenesis¹¹, but preventing the generation of neurons dying at E6 by anti-NT-3 treatment might cancel out any possible increase in cell death caused by the anti-NT-3 antibody.

As p75 is also expressed on many cells that are not dying, we investigated whether the application of exogenous NGF would increase cell death. NGF was injected either directly into one eye at E5 ($1 \mu\text{l}$, $7.5 \text{ ng } \mu\text{l}^{-1}$, $n = 5$), using the contralateral eye as a control, or was applied systemically ($2.5 \mu\text{g}$ NGF per embryo per day from E2.5 to E5.5, $n = 4$) using retinæ from untreated embryos ($n = 5$) as a control. When soluble nucleosomes were quantified in E6 embryos, no increase in cell death could be observed. Thus expression of p75 alone is not sufficient to predict NGF-mediated cell death. Other conditions must presumably be met, such as the presence of cytoplasmic interactors mediating the effects of p75, and/or the activity of the still ill-defined cell-death program in p75-expressing cells. In embryos not treated with anti-NGF antibodies, p75 is likely to mediate cell death as no other receptors have been described for this neurotrophin, and no *trkA* is expressed during this early developmental period (Fig. 1). To demonstrate this point directly, we used anti-p75 antibodies that were raised against the extracellular domain of chick p75 and which have previously been shown to block the binding of NGF to this receptor¹². Injecting these antibodies into the vitreal cavity of E4.5 embryos led to a substantial reduction in cell death, indicating that the death of retinal cells is mediated by a p75-dependent mechanism (Table 1).

These results indicate that the binding of NGF to p75 accounts for much of the early cell death during retinal development. The application of NGF has previously been reported to increase cell death in the isthmo-optic nucleus¹³, and the death of motor neurons in the facial nucleus after sectioning of the facial nerve in new-born rats¹⁴; in both cases, cell death was prevented by BDNF. Given our results, it is possible that in these previous experiments the death induced by NGF might have resulted from the activation of p75, which is known to be expressed both by neurons of the isthmo-optic nucleus and by axotomized motor neurons^{7,15}. Alternatively, exogenous NGF might prevent binding of endogenous ligands to p75 thought to help in the detection of low

TABLE 1 Cell death in E6 retina after antibody treatments

Control	Anti-NGF	Anti-p75
1.01 ± 0.11 ($n = 9$)†	$0.47 \pm 0.17^*$ ($n = 12$)	—
1.13 ± 0.34 ($n = 5$)‡	—	$0.47 \pm 0.16^*$ ($n = 11$)

Soluble nucleosomes were quantified in retinæ from control and antibody-treated embryos. For each determination, the absorbance obtained in controls was normalized to a value of 1; n , total number of embryos used in 3 separate experiments with either anti-NGF or with anti-p75 antibodies.

* $P < 0.001$ (Student's *t*-test).

† Controls were injected with PBS.

‡ Controls were injected with an irrelevant antiserum.

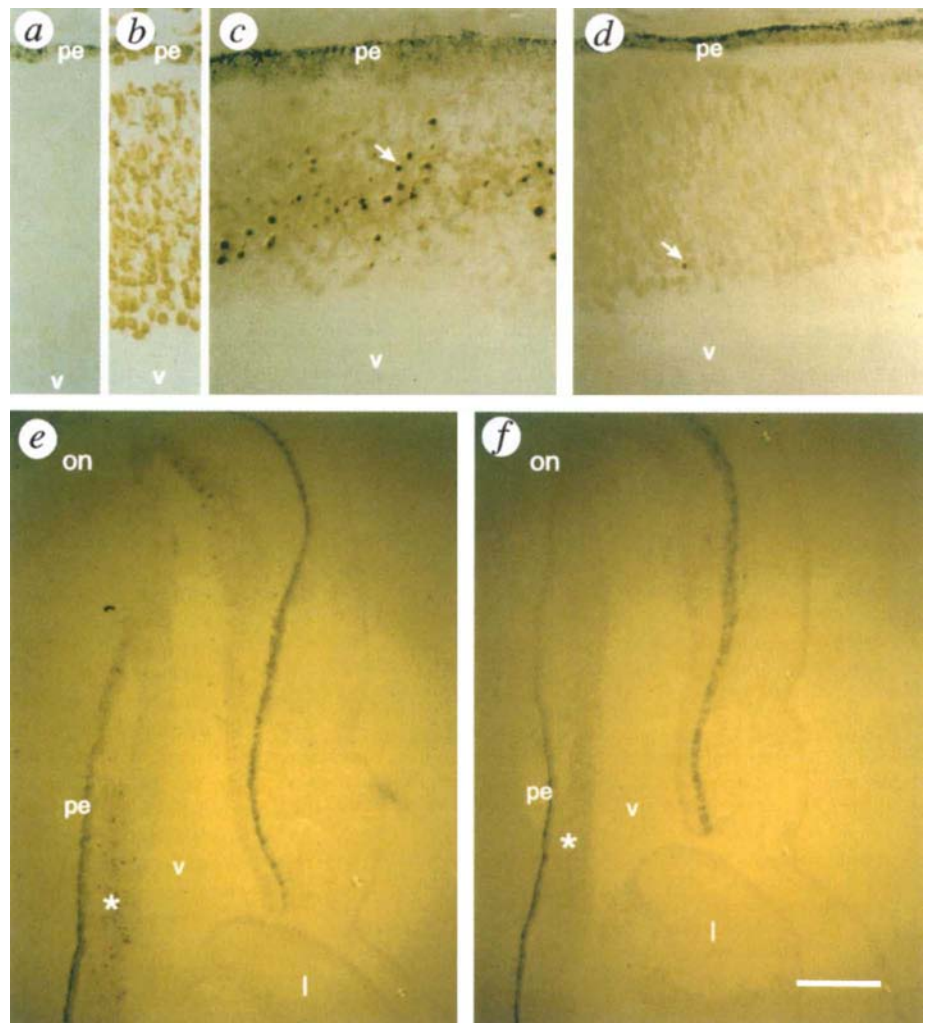


FIG. 2 NGF-antibodies reduce cell death in the E6 chick retina, as assessed by the terminal transferase reaction. *a*, Control TUNEL staining without incubation with terminal transferase. *b*, TUNEL staining on adjacent section after DNase treatment. *c*, TUNEL staining of control, untreated embryos. Note abundant labelled nuclei (arrow). *d*, The same retinal region as *c*, from an anti-NGF-treated embryo. Note the dramatic reduction in the number of stained cells. *e*, Control retina at a lower magnification showing numerous stained nuclei in the central retina. The region with the asterisk corresponds to that shown in *c*. *f*, Anti-NGF-treated retina. The region by the asterisk corresponds to that shown in *d*. Abbreviations: l, lens; on, optic nerve; pe, pigment epithelium; v, vitreous body. Scale bar, $25 \mu\text{m}$ (*a-d*), $150 \mu\text{m}$ (*e, f*). Similar results were obtained in four separate experiments.

levels of neurotrophin by the *trk* receptors. Previous *in vitro* studies^{16,17} have suggested a role for p75 in mediating cell death, but reached the conclusion that NGF blocked cell death, whereas we found that NGF caused cell death *in vivo*.

Thus, NGF seems to be unique endogenous ligand needed both to cause and prevent cell death during normal development. *In vitro*, ligand-mediated elimination of nerve cells is already well documented in the developing nervous system. In particular, the bone morphogenetic protein 4 (Bmp-4) has been shown to induce neural crest depletion in rhombomeres 3 and 5 (ref. 18). Ligand-mediated cell death might be part of a morphogenetic process, and/or contribute to the elimination of unwanted cellular phenotypes. It will be interesting to examine whether NGF participates in the elimination of the neurons that form transient structures, such as the cortical subplate. This transient cell population is involved in the formation of ocular dominance columns, and expresses p75 before its elimination^{19,20}. In the central retina, ligand-mediated cell death might create space for the axons of the retinal ganglion cells converging towards the central retina to form the optic nerve. Indeed, the distribution of dying cells, as observed on reconstruction of retinal section, seems to be related to the development of the optic fibres⁸. □

Methods

Embryos. Fertilized eggs from white-leghorn hens were incubated at 38.5 °C in 70% humidity, and were staged according to published methods²¹.

Reverse transcription-polymerase chain reaction (RT-PCR). Total RNA from E4, E6 and E14 retinae was extracted using TRIZOL reagent (Gibco BRL), and 500 µg total RNA was used for cDNA synthesis using SuperScript II RNase H⁻ reverse transcriptase (Gibco BRL), and 16 µg cDNA used for PCR amplification (NGF primers, bp 595–614 and 804–823, ref. 22; *trka* primers, 1443–1462 and 1743–1762, ref. 23; glyceraldehyde-3-phosphate dehydrogenase, 659–678 and 1186–1205, ref. 24). NGF and *trka* mRNAs were amplified for 33 cycles (59 °C for 30 s/72 °C for 1 min/95 °C for 30 s), and GAPDH mRNA for 30 cycles (65 °C for 30 s/72 °C for 1 min/95 °C for 30 s). Portions of the reactions were fractionated in 1% agarose gels. Amplification reactions in the absence of reverse transcription did not lead to the detection of any reaction product (not shown).

TUNEL staining. E6 chicken embryo heads were fixed in 4% para-formaldehyde/PBS overnight at 4 °C, incubated in 100 mM sodium phosphate, pH 7.3, containing 30% sucrose. Cryopreserved retinal sections (8 µm) were stained with the *in situ* cell-death detection kit (Boehringer Mannheim) following the manufacturer's instructions.

Quantification of cell death. Cell death was quantified by an enzyme-linked immunosorbent assay (ELISA) (Boehringer Mannheim) using a combination of antibodies recognizing histones and DNA, allowing the quantification of soluble nucleosomes in cell lysates¹⁰. Retinae of control and treated embryos were homogenized in 200 µl PBS containing 1 mM PMSF and centrifuged at 20,000g for 10 min. A portion of supernatant was used to quantify proteins by standard methods, and the rest was diluted 1:10 in the supplied buffer and processed. Absorbance values ranged between 0.1 and 0.7, and were normalized with respect to the values obtained with control retinae.

Embryo treatment with hybridoma cells. Hybridoma cells secreting anti-NGF or anti-NT-3 antibodies (both IgG1) were grown *in ovo* following the procedure previously described for quail embryos^{9,25}. Eggs were incubated for 2.5 days (until stage 14–15 of development), and a suspension of 2×10^6 hybridoma cells in 50 µl PBS was applied onto the embryos that were killed after 3.5 days. In the anti-NGF-treated embryos, 44 ± 9 µg anti-NGF immunoglobulin per g protein was measured in the retinae.

Eye injections of anti-p75. In the anti-p75 experiments, 1 µl of the CHEX antisera (diluted 1:10 in PBS) was injected into the vitreal space of E4.5 chicken eyes using a Microinjector 5242 (Eppendorf). After 12 h, 7 ± 1.7 µg immunoglobulin per g protein was detected in the eye, and nucleosomes quantified and compared with the levels determined in the contralateral, control eye. Embryos treated with an irrelevant antisera diluted 1:10 in PBS were also used as controls.

8. Cuadros, M. A. & Rios, A. *Anat. Embryol.* **178**, 543–551 (1988).
9. Rohrer, H. et al. *Development* **103**, 545–552 (1988).
10. Leist, M. et al. *J. Immunol.* **153**, 1778–1788 (1994).
11. Bovolenta, P. et al. *J. Neurosci.* **16**, 4402–4410 (1996).
12. Weskamp, G. & Reichardt, L. F. *Neuron* **6**, 649–663 (1991).
13. von Bartheld, C. S. et al. *Neuron* **12**, 639–654 (1994).
14. Sendtner, M., Holtmann, B., Kolbeck, R., Thoenen, H. & Barde, Y.-A. *Nature* **360**, 757–759 (1992).
15. Ernfrors, P., Henschen, A., Olson, L. & Persson, H. *Neuron* **2**, 1605–1613 (1989).
16. Rabizadeh, S. et al. *Science* **261**, 345–348 (1993).
17. Barrett, G. L. & Bartlett, P. F. *Proc. Natl Acad. Sci. USA* **91**, 6501–6505 (1994).
18. Graham, A., Francis-West, P., Brickell, P. & Lumsden, A. *Nature* **372**, 684–686 (1994).
19. Ghosh, A. & Shatz, C. J. *Science* **255**, 1441–1443 (1992).
20. Allendoerfer, K. L. et al. *J. Neurosci.* **14**, 1795–1811 (1994).
21. Hamburger, V. & Hamilton, H. J. *Morphol.* **88**, 49–92 (1951).
22. Ebendal, T., Larhammar, D. & Persson, H. *EMBO J.* **5**, 1483–1487 (1986).
23. Schröpel, A., von Schack, D., Dechant, G. & Barde, Y.-A. *Mol. Cell. Neurosci.* **6**, 544–556 (1995).
24. Dugaiczky, A. et al. *Biochemistry* **22**, 1605–1613 (1983).
25. Gaese, F., Kolbeck, R. & Barde, Y.-A. *Development* **120**, 1613–1619 (1994).

ACKNOWLEDGEMENTS. We thank L. R. Reichardt for providing the CHEX antiserum.

CORRESPONDENCE and requests for materials should be addressed to Y.-A.B.

Receptor-associated Mad homologues synergize as effectors of the TGF-β response

Ying Zhang, Xin-Hua Feng, Rui-Yun Wu & Rik Derynck

Departments of Growth and Development, and Anatomy, Programs in Cell Biology and Developmental Biology, University of California at San Francisco, San Francisco, California 94143-0640, USA

TRANSFORMING growth factor-β TGF-β is the prototype for a family of extracellular proteins that affect cell proliferation and tissue differentiation^{1–3}. TGF-β-related factors, including BMP-2/4, Dpp and activin, act through two types of serine/threonine kinase receptors which can form a heteromeric complex^{3,4}. However, the mechanism of signal transduction by these receptors is largely unknown. In *Drosophila*, Mad is required for signalling by Dpp⁵. We have isolated complementary DNAs for four human Mad homologues, one of which, hMAD-4, is identical to DPC-4, a candidate tumour suppressor⁶. hMAD-3 and -4 synergized to induce strong ligand-independent TGF-β-like responses. When truncated at their carboxy termini, hMAD-3 and -4 act as dominant-negative inhibitors of the normal TGF-β response. The activity of hMAD-3 and -4 was regulated by the TGF-β receptors, and hMAD-3 but not hMAD-4 was phosphorylated and associated with the ligand-bound receptor complex. These results define hMAD-3 and -4 as effectors of the TGF-β response and demonstrate a function for DPC-4/hMAD-4 as a tumour suppressor.

Several vertebrate Mad homologues have been implicated in BMP-2/4 and activin signalling. MADR1/Smad-1 is phosphorylated following coexpression with type I and type II BMP-2/4 receptors⁷, and *Xenopus* XMad-1 and XMad-2 induce BMP-2/4-like ventral or activin-like dorsal mesoderm, respectively, in *Xenopus* embryos⁸. No Mad homologues involved in TGF-β signalling have been identified. We have isolated cDNAs for four human hMAD homologues, named hMAD-1 to 4, in which hMAD-1 corresponds to MADR1/Smad-1 and hMAD-4 to the candidate tumour suppressor DPC-4 (Fig. 1). To evaluate their involvement in TGF-β signalling, we tested their ability to induce luciferase expression from a PAI-1 (plasminogen activator inhibitor-1) promoter, a common assay for TGF-β-induced gene expression⁹. Expression plasmids which expressed C-terminally FLAG-tagged hMAD-1 to -4 at similar levels (data not shown) were transfected in TGF-β-unresponsive SW480.7 cells. This cell line¹⁰ lacks expression of hMAD-4, but not hMAD-1 and -3 (data not shown), consistent with its deletion of the 18q21 locus^{6,10}.

Received 20 June; accepted 30 July 1996.

1. Oppenheim, R. W. *Annu. Rev. Neurosci.* **14**, 453–501 (1991).
2. Lewin, G. R. & Barde, Y.-A. *Annu. Rev. Neurosci.* **19**, 289–317 (1996).
3. Barbacide, M. J. *Neurobiol.* **25**, 1386–1403 (1994).
4. Blaschke, A. J., Staley, K. & Chun, J. *Development* **122**, 1165–1174 (1996).
5. Sulston, J. & Horvitz, H. R. *Dev. Biol.* **56**, 110–156 (1977).
6. Carter, B. D. et al. *Science* **272**, 542–545 (1996).
7. von Bartheld, C. S., Heuer, J. G. & Bothwell, M. J. *Comp. Neurol.* **310**, 103–129 (1991).

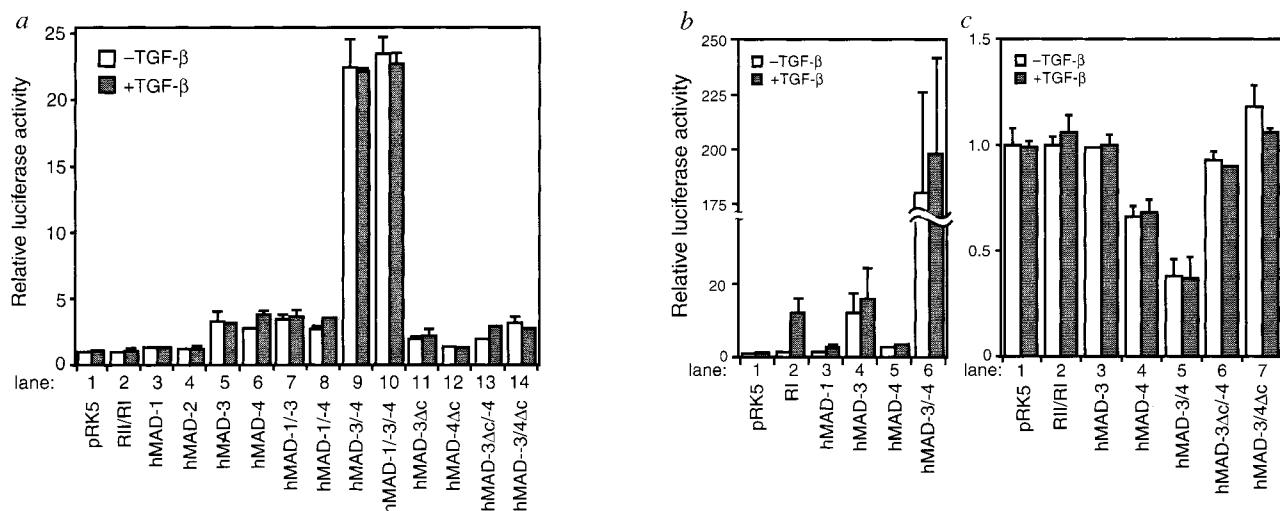


FIG. 2 hMAD-3 and hMAD-4 synergize to induce TGF- β like responses in SW480.7 (a, c) and RI-deficient Mv1LuR1B cells (b). a, Induction of luciferase expression from the PAI-1 promoter¹³ by transfected hMADs. The values are relative to control (pRK5) transfected cells in the absence of TGF- β . hMAD-3 or -4, but not hMAD-1 or -2, induce a low level of PAI-1-luciferase activity (lanes 3–6), and hMAD-3 and -4 synergize to induce a high response (lanes 9, 10). hMAD-3 Δ c and -4 Δ c are inactive (lanes 11–14). The responses are ligand-independent. b, Induction of luciferase (as in

a) in TGF- β -unresponsive R1B cells, which lack functional RI¹², by transfected hMADs. Expression of transfected RI confers TGF- β responsiveness. c, Decreased cyclin A transcription, measured by luciferase expression from pCAL2¹³, in hMAD-transfected SW480.7 cells. Relative values are calculated based on cyclin A-luciferase expression in untreated, control transfected cells (lane 1). Expression of hMAD-4, but not hMAD-3, decreases cyclin A expression. hMAD-3 and -4 synergize to induce a stronger response. hMAD-3 Δ c and -4 Δ c are inactive.

complex coprecipitated with hMAD-3, but not hMAD-4 (Fig. 5a). The receptor association of hMAD-3 but not hMAD-4 is consistent with the phosphorylation of hMAD-3, but not hMAD-4 *in vivo* (Fig. 5b). The phosphorylation of hMAD-3 on serine and less on threonine (Fig. 5c) is also consistent with the kinase specificity of the receptors. Finally, hMAD-3, and to a much lesser extent hMAD-4, was phosphorylated *in vitro* by the kinase domain of RI and only poorly by RII (Fig. 5d).

The functional interactions of hMAD-3 and -4 in TGF- β receptor signalling could be explained by their physical association. hMAD-3 and -4 displayed strong heteromeric and homomeric interactions in yeast two hybrid assays¹⁶ (data not shown). The heteromeric interactions could not be shown by coimmunoprecipitations, which might be due to the transient nature of this interaction or a required stabilization by additional component.

Previous observations suggested an involvement of hMAD-1/

MADR1/Smad-1 and the *Xenopus* homologue XMad-1 primarily in BMP-2/4 signalling^{7,8,17}, whereas Xmad-2, the equivalent of hMAD-2, may be involved in activin signalling⁸. We have now characterized hMAD-3 and hMAD-4 as effectors of TGF- β signalling. Their synergistic activity is regulated by the TGF- β receptor kinase complex, and hMAD-3, but not -4, associates with the receptor complex, consistent with the phosphorylation of hMAD-3 but not -4. Together with the cytoplasmic and nuclear localization of Mad homologues^{7,8,17} and their possible function as transcriptional regulators¹⁷, our results suggest a receptor signalling model that resembles the cytokine-induced STAT (for signal transducer and activator of transcription) responses¹⁸. Following tyrosine phosphorylation by receptor-associated JAK kinases, a phosphorylated heteromeric STAT complex is translocated into the nucleus where it functions as transcriptional activator¹⁸. Similarly, hMAD-3 may be a direct substrate for the receptor

FIG. 3 Truncated hMADs are dominant negative inhibitors of the TGF- β response. a, PAI-1-luciferase expression in SW480.7 cells (as in Fig. 2a). Overexpression of hMAD-3 Δ c or -4 Δ c inhibits the PAI-1 activity induced by low levels (0.1) of hMAD-3 and -4 (lanes 5–7). b, Dominant-negative inhibition of TGF- β -induced PAI-1 expression (100%) in Mv1Lu cells by overexpressing truncated RI or RII (lanes 1–3), or truncated hMAD-3 Δ c or -4 Δ c, but not -1 Δ c (lanes 4–7). c, Truncated hMAD-3 Δ c or -4 Δ c inhibit the antiproliferative effect of TGF- β as measured by DNA synthesis. R1B cells that lack RI are not growth inhibited (lane 1) unless transfected with RI (lane 2). Coexpression with hMAD-3 Δ c or -4 Δ c inhibits the antiproliferative response by TGF- β (lanes 3, 4). ³H-thymidine incorporation in the 60% (on average) of untransfected cells was subtracted to generate the values for only the transfected cells.

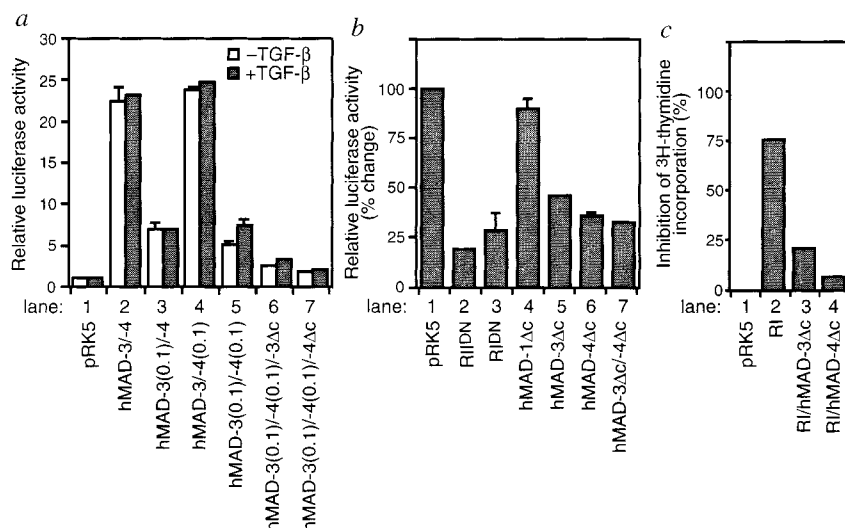


FIG. 4 hMAD-3 and -4 activity is regulated by TGF- β receptors in SW480.7 cells. PAI-1-luciferase expression is shown as in Fig. 2a. **a**, Overexpression of cytoplasmically truncated forms of RI and/or RII (lanes 3–5), which function as dominant-negative inhibitors of TGF- β receptor function, but not of related activin RII (lane 6) or Tsk7L RI (lane 7), inhibits the activity of hMAD-3 and -4. **b**, Coexpression of RII and RI enhances the activity of hMAD-3 and -4 expressed at low level (0.1). The activity of hMAD-3 and -4 (lane 4) is dramatically enhanced by coexpression of RII and RI (lane 5) but not by either receptor alone (lanes 8 and 9), or kinase-deficient receptors (RIIKR and RIIKR; lanes 10–12), or RII and Tsk7L RI (lane 13).

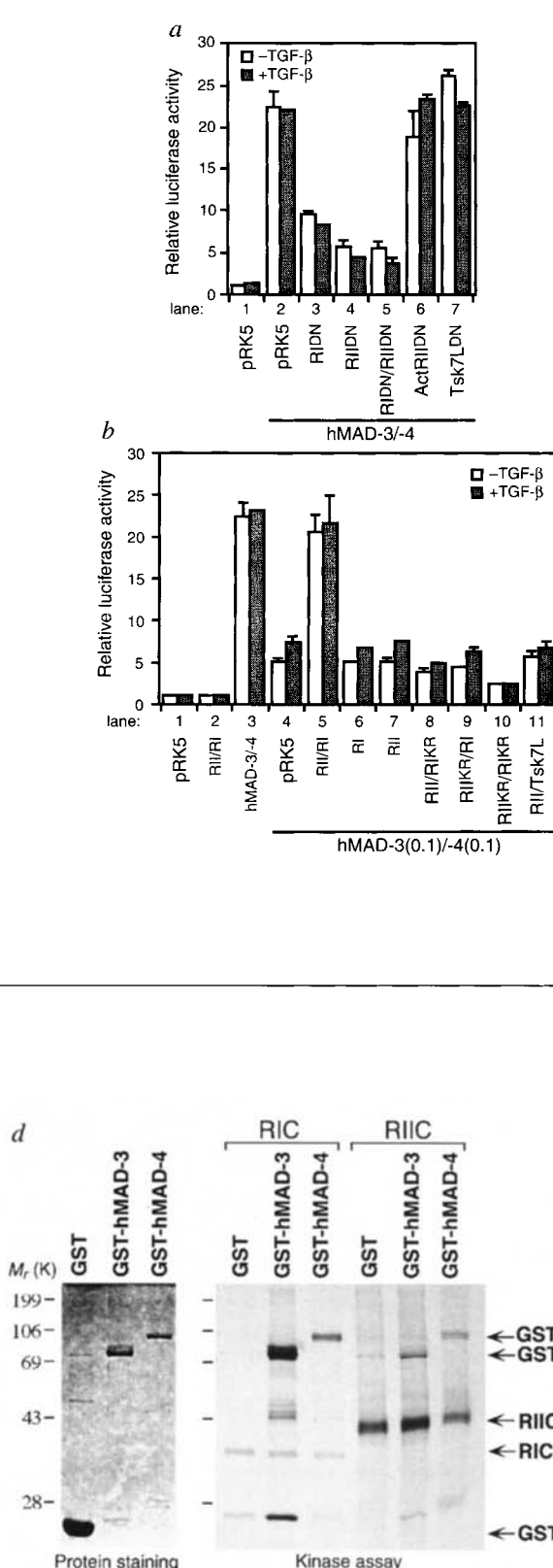
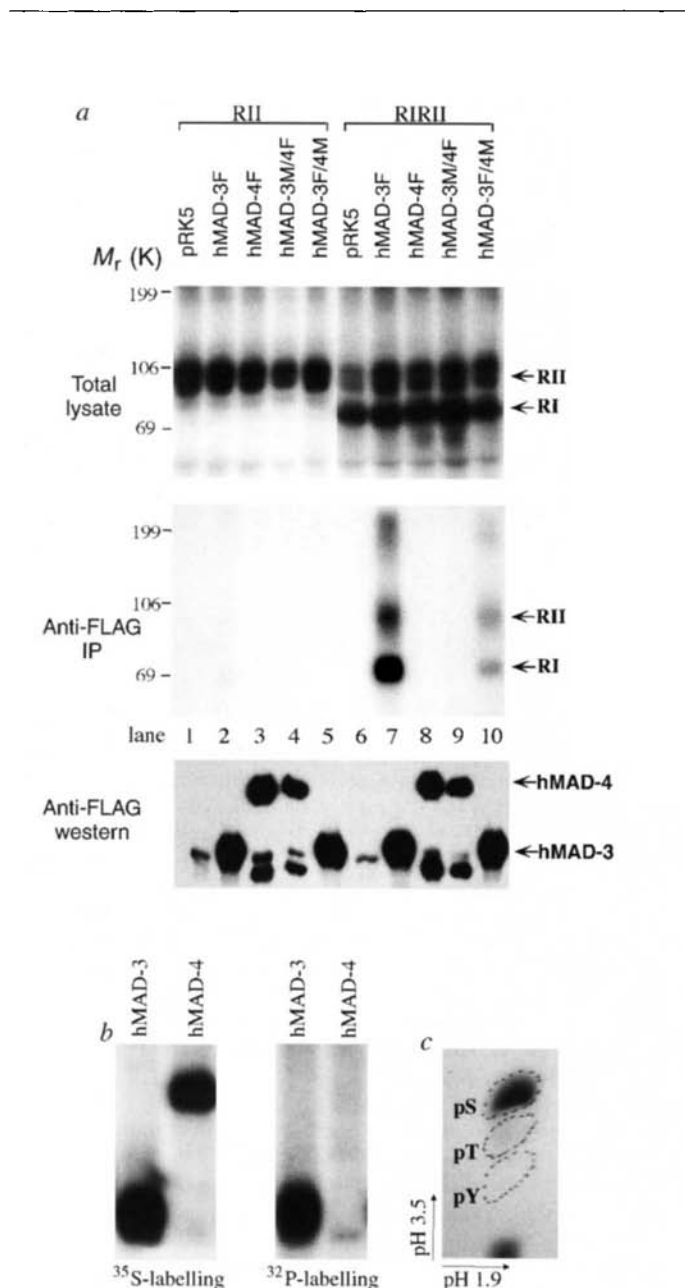


FIG. 5 Receptor association and phosphorylation of hMAD-3, but not hMAD-4. **a**, 293 cells were transfected with expression plasmids for RII or RI and RII (RIRII), and for Myc (M)- or FLAG (F)-tagged hMAD-3 and -4. ¹²⁵I-TGF- β binding and crosslinking allowed detection of receptor complexes at the cell surface. Top, pattern characteristic for RII or RII and RI. Immunoprecipitations using FLAG antibody (middle panel) showed coprecipitation of ligand-bound RI and RII with hMAD-3 (lanes 7, 10) and not hMAD-4 (lanes 8, 9), irrespective of whether hMAD-4 was coexpressed with hMAD-3 or not. hMAD-3 or -4 did not associate with RII alone (lanes 2–5). Bottom, hMAD-3

and -4 levels, as assessed by anti-FLAG western blot, in the same coimmunoprecipitation experiment. **b**, *In vivo* phosphorylation of hMAD-3 and -4. hMAD-3, but not hMAD-4, is ³²P-phosphorylated (right) even though expressed at similar levels (³⁵S-labelling, left). **c**, Phosphoamino acid analysis of *in vivo* ³²P-labelled hMAD-3. **d**, *In vitro* ³³P-phosphorylation of GST-hMAD-3 and -4 by purified cytoplasmic domains of RI (RIC) or RII (RIIC). hMAD-3, and to a lesser extent hMAD-4, are phosphorylated by RIC, but only poorly by RIIC. RIC and RIIC are autophosphorylated.

kinase, and a complex of hMAD-3 with hMAD-4 and possibly other proteins may function as a transcriptional regulator of downstream genes in the nucleus. Proteins associated with this hMAD complex may be homologues of *schnum*^{19,20}, *sal* and *sah*²¹, which in *Drosophila* act downstream from Dpp and structurally resemble transcription factors. Our results also define a role for the candidate tumour suppressor DPC-4, because its inactivation prevents TGF- β responsiveness and the resulting growth inhibition. Accordingly, SW480.7 cells lack hMAD-4 and are TGF- β -irresponsive, despite the presence of TGF- β receptors. The requirement of hMAD-3 in TGF- β signalling further predicts a function of hMAD-3 as tumour suppressor. A function of TGF- β signalling mediators as tumour suppressors is consistent with the inactivation of RII in several carcinomas²² and a role of pRB in tumour suppression and TGF- β signalling²³. The disintegration of the TGF- β signalling pathway may greatly contribute to the progression towards malignancy.

Note added in proof: hMAD-2 (91.5% identity with hMAD-3), is identical to human MADR2 (ref. 30) and JV18-1 (ref. 31) and corresponds to *Xenopus* Xmad-2 (ref. 8) and mouse Madr2 (ref. 32). In contrast with hMAD-3 Δ c or -4 Δ c (Fig. 3b), over-expression of hMAD-2 Δ c did not inhibit the TGF- β -induced PAI-1 response in Mv1Lu cells. □

Methods

hMAD cDNAs and expression plasmids. The cDNAs for hMAD-1 to -4 were isolated by combining IMAGE Consortium cDNAs²⁴, identified by BLAST searches against the data base of expressed sequence tags (dbEST), with cDNA cloning from a human placenta cDNA library. The full-size coding sequences were subcloned into pRK5F or pRK5M¹³ to express C-terminally FLAG- or Myc-tagged hMADs. Expression plasmids for the receptors and their mutants were described before^{13,25}.

Transfections, metabolic labelling, ¹²⁵I-TGF- β receptor crosslinking and immunoprecipitations. 293 cells were transfected using LipofectAMINE (Life Technologies) and metabolically ³⁵S- or ³²P-labelled overnight²⁵. Crosslinking of cell-surface TGF- β receptors²⁶ with ¹²⁵I-labelled TGF β 1 (NEN), immunoprecipitations using anti-FLAG M2 antibody¹³ (IBI, Kodak) and phosphoamino acid analysis²⁵ were done as described.

Functional TGF- β assays. Reporter assays were done as described¹³ in SW480.7 or Mv1Lu cells, or the derivative R1B and DR26 cells, transfected by LipofectAMINE or electroporation with plasmids p800luc for the PAI-1 reporter assay²⁷ or pCAL2 for the cyclin A reporter assay¹³. Total amounts of transfected DNA were the same in all assays and all values were normalized for β -galactosidase expression from the cotransfected plasmid pRK β gal¹³. DNA synthesis was assayed in transfected R1B-L17 cells by radiolabelling with ³H-thymidine (NEN) for 4 h and measuring acid-insoluble ³H-thymidine²⁸.

GST fusion proteins and *in vitro* kinase assays. Full-length hMAD-3 and -4 coding sequences were subcloned into BamHI-EcoRI cut pGEX-2T (Pharmacia) to allow expression of glutathione-S-transferase (GST)-hMAD-3 or -4 fusions in *E. coli*. Purified GST, GST-hMAD-3 or GST-hMAD-4 bound to glutathione-Sepharose²⁹ was subjected to kinase assays in the presence or absence of the cytoplasmic domain of RI or RII (purified from baculovirus-infected insect cells).

Yeast two-hybrid analysis. Full-length coding regions of hMAD-3 or -4 were cloned in-frame into the EcoRI and XhoI sites of pEG202 and pJG4-5 and interaction analysis was done as described¹⁶.

Received 15 July; accepted 13 August 1996.

1. Derynck, R. In *The Cytokine Handbook* (ed. Thomson, A. W.) 319–342 (Academic, San Diego, 1994).
2. Kingsley, D. M. *Genes Dev.* **8**, 133–146 (1994).
3. Massagué, J., Attisano, L. & Wrana, J. L. *Trends Cell Biol.* **4**, 172–177 (1994).
4. Derynck, R. *Trends Biochem. Sci.* **19**, 548–553 (1994).
5. Sekelsky, J. J., Newfeld, S. J., Raftery, L. A., Chartoff, E. H. & Gelbart, W. M. *Genetics* **139**, 1347–1358 (1995).
6. Hahn, S. A. et al. *Science* **271**, 350–353 (1996).
7. Hoodless, P. A. et al. *Cell* **85**, 489–500 (1996).
8. Graff, J. M., Bansal, A. & Melton, D. A. *Cell* **85**, 479–487 (1996).
9. Abe, M. et al. *Anal. Biochem.* **216**, 276–284 (1994).
10. Goyette, M. C. et al. *Mol. Cell. Biol.* **12**, 1387–1395 (1994).
11. Wrana, J. L. et al. *Cell* **71**, 1003–1014 (1992).
12. Weis-Garcia, F. & Massagué, J. *EMBO J.* **15**, 276–289 (1996).
13. Feng, X.-H., Filvaroff, E. H. & Derynck, R. *J. Biol. Chem.* **270**, 24237–24245 (1995).
14. Matzuk, M. M. & Bradley, A. *Biochim. Biophys. Acta* **1130**, 105–108 (1992) (erratum, 243).
15. Ebner, R. et al. *Science* **260**, 1344–1348 (1993).
16. Gyuris, J., Golemis, E., Chertkov, H. & Brent, R. *Cell* **75**, 791–803 (1993).
17. Liu, F. et al. *Nature* **381**, 620–623 (1996).
18. Ihle, J. *Nature* **377**, 591–594 (1995).
19. Aora, K. et al. *Cell* **81**, 781–790 (1995).

20. Grieder, N. C., Nellen, D., Burke, R., Basler, K. & Affolter, M. *Cell* **81**, 791–800 (1995).
21. de Celis, J. F., Barrio, R. & Kafatos, F. C. *Nature* **381**, 421–424 (1996).
22. Markowitz, S. et al. *Science* **268**, 1336–1338 (1995).
23. Laiho, M., DeCaprio, J. A., Ludlow, J. W. & Livingston, D. M. *Cell* **62**, 175–185 (1990).
24. Lennon, G. G., Auffray, C., Polymeropoulos, M. & Soares, M. B. *Genomics* **33**, 151–152 (1996).
25. Chen, R.-H., Miettinen, P. J., Maruoka, E. M., Choy, L. & Derynck, R. *Nature* **377**, 548–552 (1995).
26. Lin, H. Y., Wang, X.-F., Ng-Eaton, E., Weinberg, R. A. & Lodish, H. F. *Cell* **68**, 775–785 (1992).
27. Keeton, M. R., Curriden, S. A., van Zonneveld, A. & Loskutoff, D. J. *J. Biol. Chem.* **266**, 23048–23052 (1991).
28. Like, B. & Massagué, J. *J. Biol. Chem.* **261**, 13426–13429 (1986).
29. Smith, D. B. & Johnson, K. S. *Gene* **67**, 31–40 (1990).
30. Eppert, K. et al. *Cell* **86**, 543–552 (1996).
31. Riggan, G. J. et al. *Nature Genet.* **13**, 347–349 (1996).
32. Baker, J. C. & Harland, R. M. *Genes Dev.* **10**, 1880–1889 (1996).

ACKNOWLEDGEMENTS. This research was supported by a grant from the NCI to R.D. and an American Cancer Society postdoctoral fellowship to X.-H.F. We thank E. Stanbridge for the SW480.7 cells, Y. Cheng for discussions and assistance, and D. Stainier, L. Choy and E. Filvaroff for critically reading the manuscript.

CORRESPONDENCE and requests for materials should be addressed to R.D. (e-mail: derynck@itsa.ucsf.edu).

RGS family members: GTPase-activating proteins for heterotrimeric G-protein α -subunits

Ned Watson*, Maurine E. Linder*, Kirk M. Druey†, John H. Kehrl† & Kendall J. Blumer*

* Department of Cell Biology and Physiology, Washington University School of Medicine, 660 South Euclid Avenue, St Louis, Missouri 63110, USA
† Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

SIGNALLING pathways using heterotrimeric guanine-nucleotide-binding-proteins (G proteins) trigger physiological responses elicited by hormones, neurotransmitters and sensory stimuli^{1,2}. GTP binding activates G proteins by dissociating Ga from G $\beta\gamma$ subunits, and GTP hydrolysis by Ga subunits deactivates G proteins by allowing heterotrimers to reform. However, deactivation of G-protein signalling pathways *in vivo* can occur 10- to 100-fold faster than the rate of GTP hydrolysis of Ga subunits *in vitro*^{3–8}, suggesting that GTPase-activating proteins (GAPs) deactivate Ga subunits. Here we report that RGS^{9,10} (for regulator of G-protein signalling) proteins are GAPs for Ga subunits. RGS1, RGS4 and GAIP (for Ga-interacting protein¹⁷) bind specifically and tightly to Ga_i and Ga_o in cell membranes treated with GDP and AlF₄[–], and are GAPs for Ga_i, Ga_o and transducin α -subunits, but not for Ga_s. Thus, these RGS proteins are likely to regulate a subset of the G-protein signalling pathways in mammalian cells. Our results provide insight into the mechanisms that govern the duration and specificity of physiological responses elicited by G-protein-mediated signalling pathways.

We initially examined the ability of recombinant His-tagged RGS1 (or His-tagged Erk2 as a negative control) to bind purified G-protein subunits *in vitro*. An interaction between RGS1 and Ga_o was detected, but was inefficient as most of the Ga subunits remained unbound (Fig. 1a). No interaction between RGS1 and G $\beta\gamma$ subunits or Ga $\beta\gamma$ heterotrimers was detected (Fig. 1b). However, the activation state of Ga subunits strongly affected their binding to RGS1 (Fig. 1c). RGS1 interacted weakly with inactive (GDP-bound) or active (GTP γ S-bound) Ga_o subunits, but bound tightly to Ga_o subunits activated with GDP and AlF₄[–]. Because AlF₄[–] induces a conformation of Ga subunits resembling the transition state leading to GTP hydrolysis^{11,12}, these results suggest that RGS1 stabilizes the transition state in the GTPase

reaction of $G\alpha_o$, as occurs in interactions between Ras and its GAP¹³. Furthermore, they suggest that RGS1 interacts with the switch regions of $G\alpha$ subunits, which change conformation during the GTPase cycle^{11,12,14-16}. Switch 1 and/or 2 may be the main recognition determinants because they adopt unique conformations in the transition state, whereas switch 3 is unlikely to be recognized as it exists in a similar conformation in the activated (GTP- γ S) and transition state (GDP + AlF_4^-) structures.

We next determined whether various RGS family members interact specifically with $G\alpha$ subunits present in membranes. His-tagged RGS1, RGS4 or GAIP was incubated with bovine brain membrane fractions treated with GDP or GDP and AlF_4^- . After detergent extracts were prepared, complexes containing RGS proteins purified by Ni^{2+} -NTA chromatography were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie-blue staining and immunoblotting using antisera specific to $G\alpha$ or $G\beta$. In membranes treated with GDP, which favours the formation of inactive G-protein heterotrimers, association of a polypeptide of relative molecular mass 40,000 (M_r 40K; the size of $G\alpha_o$ and $G\alpha_i$) with RGS1, RGS4 and GAIP was inefficient or undetectable (Fig. 2a). In membranes treated with GDP and AlF_4^- , which induces the transition-state conformation and dissociates $G\alpha$ from $G\beta\gamma$ subunits, a 40K polypeptide was recovered nearly stoichiometrically with RGS1, RGS4 and GAIP, and identified by immunoblotting experiments as a mixture of $G\alpha_i$ and $G\alpha_o$ subunits (Fig. 2b, and data not shown). In contrast, an interaction between RGS proteins and $G\alpha_s$ or $G\beta\gamma$ subunits was not detected (data not shown). These results indicate that RGS1, RGS4 and GAIP bind tightly and specifically to $G\alpha_o$ and members of the G_i family of $G\alpha$ subunits in cell membranes.

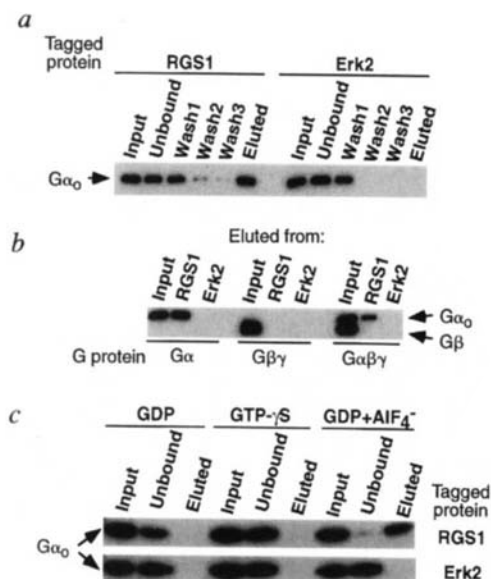


FIG. 1 Interaction of RGS1 with G-protein subunits. a, Binding of recombinant $G\alpha_o$ to His-tagged RGS1. 2% of the input, unbound and wash fractions, and 20% of the eluted fraction, were resolved by SDS-PAGE and analysed by immunoblotting using a $G\alpha$ -common antibody and enhanced chemiluminescence (ECL) detection. Binding reactions using His-tagged Erk2 were used as controls. b, Binding of $G\alpha_o$, $G\beta\gamma$ and $G\alpha\beta\gamma$ heterotrimers to His-tagged RGS1. GDP was present during the binding step and absent during the wash steps. 2% of the input fraction and 20% of the eluted fraction were analysed by immunoblotting using a mixture of $G\alpha$ -common and $G\beta$ antisera and ECL detection. In experiments using $G\alpha\beta\gamma$ heterotrimers, the fraction of $G\alpha$ subunits that bound RGS1 presumably were not incorporated into heterotrimers. c, Effects of $G\alpha_o$ activation state on RGS1 binding. $G\alpha_o$ subunits loaded with the indicated guanine nucleotides were analysed for RGS1 binding activity. The indicated nucleotides were present during the binding and wash steps. 10% of the input, unbound and eluted fractions were analysed by immunoblotting using a $G\alpha$ -common antibody and ^{125}I -labelled secondary antibodies.

To determine whether RGS proteins affect the ability of $G\alpha$ subunits to bind or hydrolyse guanine nucleotides, we examined the effects of RGS1, RGS4 and GAIP on the activity of $G\alpha_o$, $G\alpha_{i1}$, or $G\alpha_{i2}$. RGS proteins did not affect the rates of GTP-GDP exchange or steady-state hydrolysis of GTP of these $G\alpha$ subunits (data not shown).

As the steady-state rate of GTP hydrolysis of $G\alpha$ subunits is limited by the slow dissociation of GDP, rather than the fast rate of GTP hydrolysis, we therefore investigated whether RGS proteins affect the rate of GTP hydrolysis of $G\alpha$ subunits during a single catalytic turnover. $G\alpha$ subunits (or glutathione S -transferase (GST)-Rab5, a monomeric GTP-binding protein used as a control) were loaded with [γ - ^{32}P]GTP in the absence of Mg^{2+} , which prevents nucleotide hydrolysis. GTP hydrolysis was initiated by adding Mg^{2+} and unlabelled GTP with or without RGS protein. RGS1 stimulated the catalytic rate of GTP hydrolysis of $G\alpha_o$ and $G\alpha_{i1}$, but Erk2 had no effect (Fig. 3). At 5 °C, the time for half of the bound GTP to be hydrolysed ($t_{1/2}$) in the presence and absence of RGS1 were, respectively, <10 s (hydrolysis was essentially complete within 10 s, the earliest time point that could be taken) and 2–4 min, indicating at least a 10-fold increase in rate. In contrast, RGS1 did not affect the catalytic rate of GTP hydrolysis of GST-Rab5 (Fig. 3), indicating that it is GAP specific for $G\alpha$ subunits of heterotrimeric G proteins.

Similar experiments indicated that RGS1, RGS4 and GAIP function as GAPs specifically for $G\alpha_o$, $G\alpha_{i1}$, $G\alpha_{i2}$, $G\alpha_{i3}$ and transducin α -subunits, but not for $G\alpha_s$ (Fig. 3, and data not shown). Whether deactivation of $G\alpha_s$ is independent of RGS or involves other RGS family members (at least 16 have been

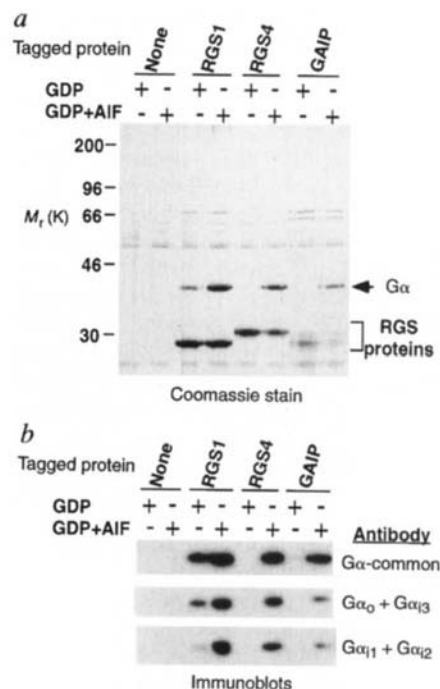
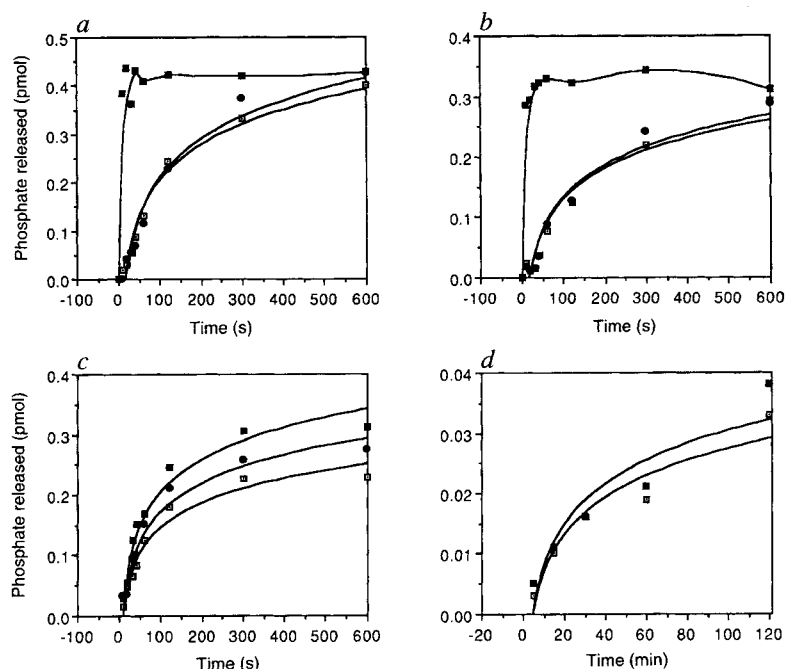


FIG. 2 Specific interaction between RGS proteins and $G\alpha$ subunits in membrane fractions. The indicated His-tagged RGS proteins were incubated with bovine brain membranes treated with GDP or GDP and AlF_4^- . After detergent extracts were prepared, complexes containing RGS proteins were purified by Ni^{2+} -NTA chromatography. a, Polypeptides bound to RGS proteins resolved by SDS-PAGE and detected by staining with Coomassie brilliant blue. b, Identification of $G\alpha$ subunits bound by RGS proteins. Immunoblotting experiments were performed with antisera specific for the indicated G-protein subunits (and $G\alpha_s$ and $G\beta$ subunits; data not shown) and ^{125}I -labelled secondary antibodies. $G\alpha$ subunits that bound to RGS proteins comigrated with recombinant, myristoylated $G\alpha_o$ and $G\alpha_i$ subunits under conditions where these species were resolved electrophoretically (data not shown).

FIG. 3 Effects of RGS1 on the catalytic rate of GTP hydrolysis of $G\alpha$ subunits: a, $G\alpha_o$; b, $G\alpha_{i1}$; c, $G\alpha_s$; and d, GST-Rab5. The kinetics of GTP hydrolysis during a single catalytic turnover by the indicated $G\alpha$ subunits and GST-Rab5 were determined at 5 °C and 37 °C, respectively, conditions necessary to quantify the effects of RGS1. The data shown are representative of at least two independent experiments. Experiments involving RGS4, GAIP, $G\alpha_{i2}$, $G\alpha_{i3}$ and transducin α -subunits were also performed at least twice. Open squares, no additions; filled circles, Erk2; filled squares, RGS1.



identified in mammalian cells^{9,10,17}) is currently unknown. However, deactivation of $G\alpha_s$ -mediated signalling in cardiomyocytes and olfactory epithelial cells occurs more rapidly than the GTPase rate of isolated $G\alpha_s$ subunits^{18–21}, suggesting that an RGS protein could be involved.

Our studies suggest a mechanism by which RGS proteins deactivate G-protein-dependent signalling pathways. RGS proteins stimulate GTP hydrolysis by $G\alpha$ subunits and then dissociate, allowing GDP-bound $G\alpha$ and $G\beta\gamma$ subunits to form inactive G-protein heterotrimers, terminating the signal. This mechanism is consistent with genetic and physiological evidence showing that RGS family members control the sensitivity of G-protein signalling pathways and the duration of cellular responses. Yeast cells lacking the RGS-related protein Sst2 are 100-fold more sensitive to mating pheromone than cells expressing Sst2 and recover slowly from pheromone-induced arrest at the G1 phase of the cell cycle; overexpression of *SST2* blunts pheromone response and promotes recovery^{22,23}. Similarly, inactivation of EGL-10, a homologue of RGS in the nematode *Caenorhabditis elegans*, enhances the function of GOA-1, a homologue of $G\alpha_o$; overexpression of *egl-10* has the opposite effect⁹.

RGS proteins are likely to have a variety of physiological functions in mammalian cells. They may provide GAP activity, which has been implicated in deactivation of G proteins that stimulate K^+ channels in atrial cardiomyocytes²⁴. RGS proteins may also function in feedback loops that promote desensitization in cells such as B lymphocytes, in which RGS expression is initially low but is induced by stimuli that activate G-protein signalling pathways^{10,25}. Finally, RGS proteins may determine whether a restricted or relatively broad spectrum of physiological responses is elicited by a particular G-protein signalling pathway². A restricted set of responses may occur in cells constitutively expressing high levels of RGS proteins, because the relatively low level and short lifetime of $G\alpha$ -GTP would allow signalling mainly through relatively abundant, high-affinity effectors. Conversely, a broader set of responses may occur in cells expressing low levels of RGS proteins, because the higher level and longer lifetime of $G\alpha$ -GTP would allow signalling through effectors of high and low affinity or abundance. Controlling the expression of RGS family members may therefore provide a versatile means of regulating G-protein signalling pathways for a variety of physiological purposes.

Note added in proof: Results similar to ours are reported by

Berman *et al.* (*Cell* **86**, 445–452; 1996) and T. W. Hunt, T. A. Fields, P. J. Casey and E. G. Peralta (*Nature*, this issue). □

Methods

Protein expression and purification. Bacterial expression plasmids for RGS proteins were constructed as follows. Vent polymerase and the entire coding regions of published RGS cDNAs^{10,17} were used to generate PCR fragments flanked by *XhoI*/*Bam*HI sites (human RGS1, human GAIP) or *NdeI*/*XhoI* sites (human RGS4) that were subcloned into the corresponding sites of the His-tag fusion vectors pET14b (RGS1) or pET15b (RGS4, GAIP) (Novagen). The GAIP cDNA was provided by L. DeVries¹⁷, and a His-tagged Erk2 expression plasmid by G. Johnson. Plasmids were transformed into the bacterial strain BL21(DE3)-pLysS (RGS1, GAIP, Erk2) or BL21(DE3) (RGS4) and induced with 1 mM IPTG. His-tagged proteins were purified by chromatography on Ni^{2+} -NTA agarose beads (Qiagen), dialysed versus HED buffer (50 mM Na-HEPES, pH 8.0, 1 mM EDTA, 1 mM DTT, 10% glycerol) and stored at –70 °C. Recombinant, myristoylated $G\alpha_{i1}$, $G\alpha_{i2}$, $G\alpha_{i3}$, $G\alpha_{o}$ ²⁶ and recombinant $G\alpha_s$ subunits²⁷ were purified from *Escherichia coli*. Purified transducin α -subunits were provided by H. Hamm. $G\beta\gamma$ subunits were purified from bovine brain²⁸. A GST-Rab5 expression plasmid was provided by P. Stahl, introduced into the bacterial strain BL21(DE3)pLysS, and induced with 1 mM IPTG. GST-Rab5 was purified by chromatography on glutathione-agarose beads.

RGS-G-protein binding assays. Binding of RGS1 to G-protein subunits required G-protein subunits to be incubated for 30 min at 5 °C with His-tagged RGS1 (1 μ g) or Erk2 (3 μ g) in 200 μ l buffer A (50 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 1 mM $MgSO_4$, 20 mM imidazole, 0.025% polyoxyethylene 10-lauryl ether ($C_{12}E_{10}$), 10 mM β -mercaptoethanol, 10% glycerol) containing 10 μ M GDP. Amounts used: $G\alpha_o$, 5 μ g; $G\beta\gamma$, 6.8 μ g; $G\alpha\beta\gamma$ heterotrimers were formed by incubating 6.8 μ g $G\beta\gamma$ and 5 μ g GDP-loaded $G\alpha_o$ subunits. Ni^{2+} -NTA agarose beads were added, incubated for 30 min at 5 °C, washed 3 times with buffer A, and eluted at 100 °C with Laemmli sample buffer. Immunoblotting experiments used antisera P960 ($G\alpha$ -common), B600 ($G\beta$ -common) and ¹²⁵I- or peroxidase-labelled secondary antibodies²⁹. Binding of RGS1 to $G\alpha_o$ loaded with GDP, GTP- γ S, or GDP and AlF_4^- required $G\alpha_o$ (5 μ g) to be incubated for 30 min at 20 °C in the presence of GDP (10 μ M), GTP- γ S (10 μ M), or GDP (10 μ M) and AlF_4^- (30 μ M)²⁷. Nucleotide-bound $G\alpha_o$ (5 μ g) was incubated with RGS1 (1 μ g) as above, except that GDP (1 μ M), GTP- γ S (1 μ M) or GDP (1 μ M) and AlF_4^- (30 μ M) was present during the binding and wash steps. For binding of His-tagged RGS proteins to G proteins in bovine brain membranes, bovine brain membranes (1 mg protein) in buffer B (20 mM Na-HEPES, pH 8.0, 380 mM NaCl, 3 mM DTT, 6 mM $MgCl_2$) containing GDP (10 μ M) or GDP (10 μ M) and AlF_4^- (30 μ M) were incubated 30 min at 5 °C with His-tagged RGS proteins (10 μ g). Membranes were solubilized with 1% cholate, and detergent-soluble extracts obtained after centrifugation at 100,000g were added to Ni^{2+} -NTA agarose beads, washed 5 times with buffer B (supplemented with 20 mM imidazole, 0.1% $C_{12}E_{10}$, 1 μ M GDP, or 1 μ M GDP and 30 μ M AlF_4^-) and eluted with 300 mM imidazole. Eluted proteins resolved by SDS-PAGE were detected by staining with Coomassie brilliant blue. The identities of G-protein subunits bound to RGS proteins were

determined by immunoblotting with antisera P960 ($G\alpha$ -common), B087 ($G\alpha_1 + G\alpha_2$), 584 ($G\alpha_3$), C260 ($G\alpha_4 + G\alpha_5$), B600 ($G\beta$ -common)²⁹ and [¹²⁵I]-labelled secondary antibodies.

Guanine-nucleotide exchange and GTPase assays. The kinetics of guanine-nucleotide exchange, steady-state GTP hydrolysis, and catalytic rates of GTP hydrolysis of $G\alpha$ subunits and GST-Rab5 were determined^{27,30}. For single turnover GTPase assays, $G\alpha$ subunits or GST-Rab5 (50–80 pmol) were incubated (15 min at 20 °C for $G\alpha_3$, and 30 °C for other $G\alpha$ subunits; 30 min at 37 °C for GST-Rab5) with [³²P]GTP (0.1 μ M; 20,000 to 30,000 c.p.m. pmol⁻¹) in 800 μ l of buffer C (50 mM Na-HEPES, pH 8.0, 5 mM EDTA, 1 mM DTT, 0.1% C₁₂E₁₀). The stoichiometries of GTP binding of $G\alpha$ subunits and GTP-Rab5 were 15–20% and 1%, respectively. GTP hydrolysis was initiated (at 5 °C for $G\alpha$ subunits; 37 °C for GST-Rab5) by adding MgSO₄ (10 mM final concentration) and unlabelled GTP (100 μ M final concentration) with or without purified RGS proteins (80 pmol) or Erk2 (80 pmol). Portions (25 μ l) removed at the indicated times were processed as described²⁷, and analysed for inorganic ³²P content by liquid scintillation spectrometry.

Received 17 July; accepted 16 August 1996.

1. Gilman, A. G. *Annu. Rev. Biochem.* **56**, 615–649 (1987).
2. Neer, E. J. *Cell* **80**, 249–257 (1995).
3. Breitwieser, G. E. & Szabo, G. J. *Gen. Physiol.* **91**, 469–493 (1988).
4. Dratz, E. A., Lewis, J. W., Schaechter, L. E., Parker, K. R. & Kilger, D. S. *Biochem. Biophys. Res. Commun.* **146**, 379–386 (1987).
5. Vuong, T. M. & Chabre, M. *Nature* **346**, 71–74 (1990).
6. Vuong, T. M. & Chabre, M. *Proc. Natl Acad. Sci. USA* **88**, 9813–9817 (1991).
7. Arshavsky, V., Gray, K. M. & Bownds, M. D. *J. Biol. Chem.* **266**, 18530–18537 (1991).
8. Arshavsky, V. & Bownds, M. D. *Nature* **357**, 416–417 (1992).
9. Koelle, M. R. & Horvitz, H. R. *Cell* **84**, 115–125 (1996).

10. Druey, K. M., Blumer, K. J., Kang, V. H. & Kehrl, J. H. *Nature* **379**, 742–746 (1996).
11. Sondek, J., Lambright, D. G., Noel, J. P., Hamm, H. E. & Sigler, P. B. *Nature* **372**, 276–279 (1994).
12. Kleuss, C., Raw, A. S., Lee, E., Sprang, S. R. & Gilman, A. G. *Proc. Natl Acad. Sci. USA* **91**, 9828–9831 (1994).
13. Mittal, R., Ahmadian, M. R., Goody, R. S. & Wittinghofer, A. *Science* **273**, 115–117 (1996).
14. Noel, J. P., Hamm, H. E. & Sigler, P. B. *Nature* **366**, 654–663 (1993).
15. Lambright, D. G., Noel, J. P., Hamm, H. E. & Sigler, P. B. *Nature* **369**, 621–628 (1994).
16. Coleman, D. E. *et al. Science* **265**, 1405–1412 (1994).
17. DeVries, L., Mousli, M., Wurmser, A. & Farquhar, M. G. *Proc. Natl Acad. Sci. USA* **92**, 11916–11920 (1995).
18. Yatani, A. & Brown, A. M. *Science* **245**, 71–74 (1989).
19. Breer, H. & Boekhoff, I. *Curr. Opin. Neurobiol.* **2**, 439–443 (1992).
20. Schleicher, S., Boekhoff, I., Arriza, J., Lefkowitz, R. J. & Breer, H. *Proc. Natl Acad. Sci. USA* **90**, 1420–1424 (1993).
21. Boekhoff, I. *et al. J. Biol. Chem.* **269**, 37–40 (1994).
22. Chan, R. K. & Otte, C. A. *Mol. Cell. Biol.* **2**, 21–29 (1982).
23. Dohliman, H. G., Song, J., Ma, D., Courchesne, W. E. & Thorne, J. *Mol. Cell. Biol.* **16**, 5194–5209 (1996).
24. Shui, Z., Boyett, M. R., Zang, W. J., Haga, T. & Kameyama, K. *J. Physiol. (Lond.)* **487**, 359–366 (1995).
25. Hong, J. X., Wilson, G. L., Fox, C. H. & Kehrl, J. H. *J. Immunol.* **150**, 3895–3904 (1993).
26. Linder, M. E. *et al. J. Biol. Chem.* **266**, 4654–4659 (1991).
27. Linder, M. E., Ewald, D. A., Miller, R. J. & Gilman, A. G. *J. Biol. Chem.* **265**, 8243–8251 (1990).
28. Stenweis, P. C. & Robshaw, J. D. *J. Biol. Chem.* **259**, 13806–13813 (1984).
29. Linder, M. E. *et al. Proc. Natl Acad. Sci. USA* **90**, 3675–3679 (1993).
30. Pan, J. Y., Sanford, J. C. & Wessling, R. M. *J. Biol. Chem.* **271**, 1322–1328 (1996).

ACKNOWLEDGEMENTS. We thank L. DeVries, G. Johnson and P. Stahl for providing cDNAs or plasmids, H. Hamm for purified transducin subunits, and K. Moulder for assistance during initial phases of this work. This work was supported in part by grants from the NIH (K.J.B. and M.E.L.).

CORRESPONDENCE and requests for materials should be addressed to K.J.B. (e-mail: kblumer@cellbio.wustl.edu).

RGS10 is a selective activator of $G\alpha_i$ GTPase activity

Timothy W. Hunt*, Timothy A. Fields†, Patrick J. Casey† & Ernest G. Peralta*

* Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

† Departments of Molecular Cancer Biology and Biochemistry, Duke University Medical School, Durham, North Carolina 27710, USA

POLYPEPTIDES that define a protein family termed RGS (for regulators of G-protein signalling) are encoded by the *SST2* gene of the yeast *Saccharomyces cerevisiae*, the *EGL-10* gene of the nematode *Caenorhabditis elegans*, and several related mammalian genes. Genetic studies in invertebrates and mammalian cell-transfection experiments indicate that RGS proteins negatively regulate signalling pathways involving seven transmembrane receptors and heterotrimeric G proteins^{1–3}. However, the biochemical mechanism by which RGS proteins control these pathways is unknown. Here we report the characterization of human RGS10, a member of this protein family. Co-immunoprecipitation studies demonstrate that RGS10 associates specifically with the activated forms of two related G-protein subunits, $G\alpha_{13}$ and $G\alpha_2$, but fails to interact with the structurally and functionally distinct $G\alpha_i$ subunit. *In vitro* assays with purified proteins indicate that RGS10 increases potently and selectively the GTP hydrolytic activity of several members of the $G\alpha_i$ family, including $G\alpha_{13}$, $G\alpha_2$ and $G\alpha_0$. These results demonstrate that RGS proteins can attenuate signalling pathways involving heterotrimeric G proteins by serving as GTPase-activating proteins for specific types of $G\alpha$ subunits.

To identify proteins that interact with the GTP-bound form of $G\alpha_{13}$, we used a yeast interaction cloning approach. Proteins encoded by a human HeLa-cell complementary DNA library were screened for their ability to interact with a mutationally activated form of $G\alpha_{13}$; the glutamine-to-leucine mutation at residue 204 (Q204L) severely impairs the intrinsic GTPase activity

of $G\alpha_{13}$ and results in a constitutive increase in the abundance of the GTP-bound form of the $G\alpha_{13}$ Q204L subunit⁴. Six novel cDNA clones were isolated from the HeLa library, one of which was identical to a cDNA present within the database generated by the Washington University–Merck expressed-sequence-tag human genome sequencing project (GenBank accession number H87415). This cDNA clone contained an open reading frame of 519 base pairs which began nine codons downstream from an in-frame stop codon. Comparison of the deduced 173-amino-acid protein with the GenBank database revealed substantial similarities with several members of the RGS family (Fig. 1). In particular, the protein encoded by the HeLa-derived cDNA clone is ~90% identical to a partial rat sequence previously termed RGS10 (ref. 2), so we tentatively refer to the $G\alpha_{13}$ -interacting protein identified by our study as human RGS10. Like all members of the family, RGS10 contains a 120-amino-acid core domain that is strongly conserved with the yeast Sst2 protein. Northern blot analysis revealed a ~900-nucleotide RGS10 transcript in messenger RNA prepared from total mouse brain, as well as human embryonic kidney 293 cells and HeLa cells (data not shown).

To investigate whether the interaction between RGS10 and $G\alpha_{13}$, originally observed through the transcriptional activation assay in yeast, could also be detected in mammalian cells, we performed co-immunoprecipitation experiments using cells transfected with RGS10 and various forms of $G\alpha$ subunits. To facilitate immunoprecipitation in these experiments, a 10-residue epitope derived from the human Myc protein was introduced immediately after the initiation methionine of RGS10. Myc-RGS10 was coexpressed with the wild-type or the mutationally activated forms of $G\alpha_{13}$, $G\alpha_2$ and $G\alpha_3$ by transient transfection of human embryonic kidney 293 cells. Proteins were immunoprecipitated from total cell lysates with an anti-Myc monoclonal antibody, and were analysed by immunoblotting with polyclonal antibodies specific for each of the coexpressed $G\alpha$ subunits. The Myc-RGS10 protein co-immunoprecipitated the activated $G\alpha_{13}$ Q204L protein, but failed to interact with the wild-type (presumably GDP-bound) form of $G\alpha_{13}$ (Fig. 2). Similarly, activated $G\alpha_2$ Q204L, but not wild-type $G\alpha_2$, co-immunoprecipitated with the Myc-RGS10 protein. In contrast, the Myc-RGS10 protein failed to co-immunoprecipitate

with either the mutationally activated or wild-type forms of G_{α_s} , a more distantly related and functionally distinct class of G_{α} protein.

Genetic studies in *S. cerevisiae* and *C. elegans*, as well as transient expression experiments in mammalian cells, indicate that members of the RGS family negatively regulate signalling pathways that are mediated by heterotrimeric G proteins¹⁻³. However, there is no direct biochemical evidence to suggest how this attenuation is achieved. Because the activated state of the G protein is the GTP-bound form, and this is the form that interacts preferentially with RGS10, we considered the possibility that an interaction with the RGS might accelerate GTP hydrolysis by the G_{α} subunit. Such a GTPase-activating protein (GAP) activity of an RGS would be consistent with a role as a negative regulator of signalling through heterotrimeric G proteins. Activity of this sort is not without precedent, as studies of the Ras family of monomeric G proteins have led to the identification of such GAPs⁵.

To investigate whether RGS proteins may serve as GAPs for activated G_{α} subunits, the ability of RGS10 to affect the GTP hydrolysis rate of several G_{α} subunits was assayed *in vitro*. A derivative of RGS10 with an amino-terminal six-histidine tag was expressed in *Escherichia coli* and purified by Ni^{2+} -resin chromatography. Purified $G_{\alpha_{13}}$, G_{α_0} , G_{α_2} and G_{α_s} proteins were prebound to [γ -³²P]GTP, and the release of the γ -phosphate was measured in the presence or absence of purified 6 $\times$ His-RGS10. Incubation with either 50 nM or 100 nM 6 $\times$ His-RGS10 dramatically increased the extent of GTP hydrolysis by $G_{\alpha_{13}}$, G_{α_0} and G_{α_2} (Fig. 3a). In contrast, neither concentration of 6 $\times$ His-RGS10 had any effect on GTP hydrolysis by purified G_{α_s} subunits. To investigate further the ability of 6 $\times$ His-RGS10 to accelerate GTP hydrolysis by G_{α} subunits, a time-course analysis was conducted using purified $G_{\alpha_{13}}$ and G_{α_s} (Fig. 3b, c). Addition of 100 nM 6 $\times$ His-RGS10 caused a >10-fold enhancement of the GTPase

RGS10	1	M....qSeLf	fadihdsgds	sssshqSlks	takwaa....	SLEn
RGS1	48	L....eSgM.	kss	kskdvLSaaE	vngfwsQ....	SLEK
GAIP	67	L....qp.L.	psc	evcatpSpeE	vqswaQ....	SFDK
Sst2p	374	irapngStid	ldftlrgrmt	vrdektdld	segfSqdml	sssnlnkLDy
Consensus		L---S-L-	---S--E	---W-Q---	-----SL--	
RGS10	37	LLeDPeGvkr	FRFLkKEFS	EENvLFWLAC	EDF..Kkmgd	ktmgqek...
RGS1	76	LLangtGgrv	FysFLkEFS	EENieFWLAC	EDY..K.kte	scdlpckK...
GAIP	94	LMnsPaGrsv	FRAFLrTEYS	EENnLFWLAC	EEL..Kaan	qhvdeK...
Sst2p	424	vLtdPgmryl	FRhLeKELC	vENldvFiei	krFLkRmtil	kkldidsKhcd
Consensus		LL--P-G---	FR-FL--EFS	EEN--FWLAC	EDF--K----	-----K---
RGS10	82			AkeLYmtFls	skassQVNVE	
RGS1	120			AeeLYkaFVh	sdaakQINID	
GAIP	139			ArllYedYVs	ilspkeVsID	
Sst2p	474	kksntstskn	nivktidsal	mkganeclen	AyhLYssYIm	igspyQINIH
Consensus		-----	-----	A--IY---V-	-----QVNID	
RGS10	102	ggsRln.eKi	LeePhplmFq	kIQdqIFn..	LMkyDSYSRF	LkSdlFLkhk
RGS1	140	frtRestakK	ikaPtptcFD	eaQkvIYt..	LMekDSYpRF	LkSdiYInl.
GAIP	159	srVReginkK	MqePsahtFD	daQlqIYt..	LMhrDSYpRF	LsSptYral.
Sst2p	524	hnlRgnisdi	MlhPhsplSE	hftptnlYdps	pasaESAass	isSt.....
Consensus		--R-----K-	M--P-----FD	--Q--IY---	LM--DSY-RF	L-S--Y----
RGS10	149	rteeeedLp	Dagtaakras	riynt		
RGS1	187	In	DlqanSlk..			
GAIP	206	Ll	qgpgSssea			
Sst2p	568	eadtLg	EpevSlkps	knlnsn		
Consensus		-----L-	D-----S----	-----		

FIG. 1 Alignment of the deduced amino-acid sequences of human RGS10 and other selected proteins reveals the conserved RGS core domain. The dots represent single-residue gaps in the sequences. The consensus sequence is denoted by upper-case letters and corresponds to residues present in at least two of the four RGS proteins. The RGS1 and GAIP clones were isolated from human B-lymphocyte and HeLa cDNA, respectively, the SST2 gene from the budding yeast *S. cerevisiae*^{1,22,23}. The N termini of the RGS1, GAIP and Sst2 proteins are truncated to the beginning of their similarity with RGS10. Numbers indicate the position of the first amino acid of each line.

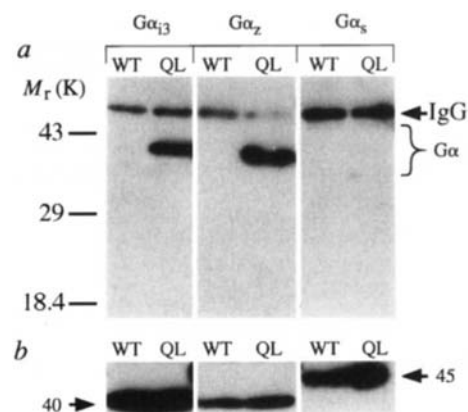


FIG. 2 RGS10 coimmunoprecipitates selectively with the activated forms of the $G_{\alpha_{13}}$ and G_{α_2} subunits, but not G_{α_s} , in mammalian cells. The wild-type (WT) and activated, GTPase-deficient mutants (QL) of $G_{\alpha_{13}}$, G_{α_2} and G_{α_s} were coexpressed with a Myc-epitope-tagged derivative of RGS10 by transient transfection of human 293 cells. *a*, Proteins were immunoprecipitated with an anti-Myc monoclonal antibody and analysed by immunoblotting with polyclonal antisera specific for the coexpressed G_{α} subunits. The positions of G_{α} subunits and immunoglobulin heavy chain (IgG) protein are indicated. *b*, To determine the relative expression levels of the transfected G_{α} subunits used in *a*, 1/50 of each cell lysate was directly subjected to immunoblot analysis with the appropriate G_{α} -specific antisera. The M_r of the expressed $G_{\alpha_{13}}$ and G_{α_2} proteins is 40K, and of G_{α_s} is 45K. Similar results were obtained in two additional experiments.

activity by $G_{\alpha_{13}}$, but G_{α_s} was again unaffected by addition of the RGS protein. We also assessed whether RGS10 could influence nucleotide exchange by a G_{α_i} subunit, namely $G_{\alpha_{13}}$; this analysis indicated that the RGS protein did not alter nucleotide exchange by $G_{\alpha_{13}}$ (data not shown). Taken together, these *in vitro* assays demonstrate that RGS10 functions as a potent GAP for several members of the G_{α_i} family, but not for the structurally distinct G_{α_s} protein.

The ability of the RGS10 protein to co-immunoprecipitate selectively with the activated forms of certain G_{α} subunits, and act as a GAP for the wild-type proteins *in vitro*, strongly suggests that RGS10, and perhaps other members of the RGS family, negatively regulate signalling pathways by reducing the lifetime of the active, GTP-bound form of G_{α_i} family subunits. Because signalling downstream of G_{α_i} -containing heterotrimers often involves regulatory interactions between effectors and $G\beta\gamma$ subunits rather than G_{α_i} -GTP, it is possible that distinct GAPs such as RGS10 are necessary to terminate such signals efficiently by promoting formation of G_{α_i} -GDP, which can reassociate with $G\beta\gamma$ (refs 6–10). The previous examples of GAP activity associated with molecules that interact with G proteins involve family members (G_{α_i} and G_{α_q}) in which the effector is regulated directly by the GTP-bound G_{α} subunit, and it is the effector itself that possesses the GAP activity^{11,12}. In such cases, a distinct RGS-like protein may not be necessary to facilitate desensitization of the G-protein signal because the effector directly influences the activation state of the G_{α} subunit. Alternatively, given the large number of RGS proteins that have been identified, it is possible that particular family members may act as GAPs that are selective for each of the various classes of G_{α} subunits.

The experiments described here do not address whether RGS10 may function as an effector or adaptor in addition to its role as a GAP for the G_{α_i} class of G-protein subunits. Nonetheless, the potent stimulatory effect of RGS10 on the GTPase activity of specific G_{α} subunits suggests that the GAP activities of RGS proteins constitute an important target for biological regulation of G-protein signalling. □

Methods

Clones. The RGS10 clone was isolated using a yeast two-hybrid cloning screen¹³. DNA encoding amino acids 26–354 of the activated Q204L mutant of

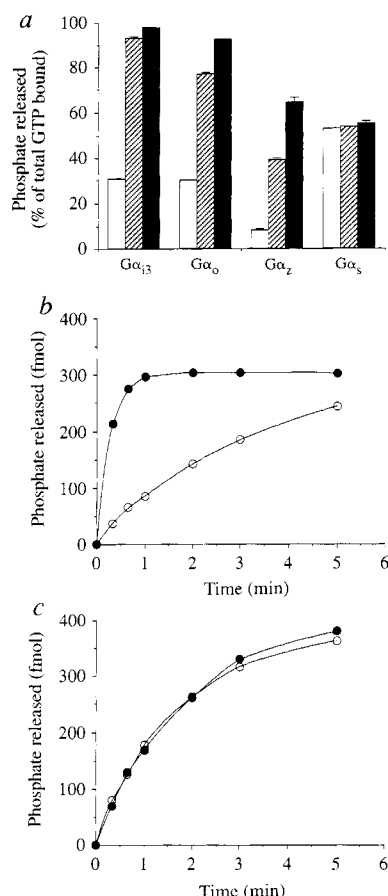


FIG. 3 RGS10 is a selective GTPase-activating protein for G α_i subunits. **a**, Purified G α subunits were loaded with [γ - 32 P]GTP and incubated without (white bars) or with 50 nM (hatched bars) or 100 nM (black bars) 6 $\times$ His-RGS10. The concentration of G α -GTP in each experiment was 12, 11, 6 or 15 nM for G α_{i3} , G α_{i2} , G α_{i1} and G α_s , respectively. GTPase assays were performed at 4 $^{\circ}$ C for 10 min (G α_{i3}), 3 min (G α_{i2}), 1 min (G α_{i1}), or 45 s (G α_s). Data shown represent the mean and s.d. of three separate experiments. **b**, Purified 6 $\times$ His-RGS10 potentially enhances the rate of GTP hydrolysis by purified G α_{i3} . G α_{i3} (10 nM) was loaded with [γ - 32 P]GTP and incubated without (filled circles) or with (open circles) 100 nM 6 $\times$ His-RGS10 at 4 $^{\circ}$ C. Portions were withdrawn at the times indicated, and GTP hydrolysis was determined by measuring the release of phosphate. **c**, Purified 6 $\times$ His-RGS10 does not alter the GTPase activity of purified G α_s . G α_s (12 nM) was loaded with [γ - 32 P]GTP and incubated without (filled circles) or with (open circles) 100 nM 6 $\times$ His-RGS10 at 4 $^{\circ}$ C. Portions were withdrawn at the times indicated, and GTP hydrolysis was determined by measuring phosphate release. Data shown are from a single experiment which is representative of at least two such experiments.

rat G α_{i3} was cloned into the pEG202 'bait' vector and cotransformed with $\sim$ 300,000 HeLa cDNA clones contained within the pJG4-5 'prey' vector using yeast strain EGY48 (refs 14, 15). Plasmid DNA recovered from 49 positive clones was partially sequenced, and comprised six independent cDNAs. cDNA inserts encoding the RGS10 protein were subcloned and sequenced on both strands by the dideoxynucleotide chain termination method. The proposed initiator methionine for the full-length RGS10 reading frame was preceded by an in-frame stop codon; this clone also contained 216 bp of 3' untranslated sequence and a poly(A) sequence.

Immunoprecipitation studies. A Myc epitope tag (MEQKLISEEDL) was fused to the 5' end of a cDNA fragment encoding amino acids 10–173 of RGS10 by site-directed mutagenesis¹⁶. All G α subunits and Myc-RGS10 were inserted into the mammalian expression vector pcDNA3 (Invitrogen) and used to transiently transfect human 293 cells¹³. The activating mutations of G α_{i3} and G α_{i2} (Q204L) or G α_{i1} (Q227L) were made⁴. Transfected cells were lysed in RIPA buffer (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, protease inhibitors and 6 mM MgCl₂) 36–48 h after transfection. Insoluble material was pelleted by centrifugation of 8,000g for 5 min. Supernatants were precleared by incubation with protein A-Sepharose beads (Pharmacia) for 1 h at 4 $^{\circ}$ C. Cleared supernatants were incubated with 3 μ g 9E10 anti-Myc monoclonal antibody overnight at 4 $^{\circ}$ C. Immunocomplexes were bound to protein A-Sepharose beads for 4 h at 4 $^{\circ}$ C, washed 3 times with 1 ml lysis buffer,

analysed by immunoblotting with rabbit polyclonal anti-G α_{i3} antisera C-10, anti-G α_{i2} antisera p961 or anti-G α_{i1} antisera RM/1, and developed using enhanced chemiluminescence.

GTPase studies. To purify the RGS10 protein for *in vitro* GTPase studies, a six-histidine tag was added to the N terminus of amino acids 10–173 of RGS10 by PCR amplification of the cDNA and cloning into the pRSETA vector (Invitrogen). The 6 $\times$ His-RGS10 protein was expressed in *E. coli* strain BL21 DE3(pLysS) and purified to >90% homogeneity (as judged on Coomassie-blue-stained protein gels) using Ni²⁺ chromatography (Invitrogen). Purified G α subunits^{17–20} were loaded with GTP by incubating each protein in 60 μ l of 50 mM HEPES, pH 7.6, 2 mM EDTA, 1 mM DTT, 0.0125% lubrol, and 5 μ M [γ - 32 P]GTP ($\sim$ 50,000 c.p.m. pmol⁻¹); the stoichiometry of loading was 25–40%. Magnesium was excluded from all loading reactions except for the G α_{i2} reaction (1.5 μ M free Mg²⁺) to slow hydrolysis. Free nucleotide was removed from the GTP-bound G proteins by centrifugation (200g for 3 min at 4 $^{\circ}$ C) through 1 ml G50 Sephadex (Pharmacia) that was equilibrated in 50 mM HEPES, pH 7.6, 1 mM EDTA and 1 mM DTT. The GTPase assay solution contained 10 μ l GTP-bound G protein and 6 $\times$ His-RGS10 in 20 μ l of 50 mM HEPES, pH 7.6, 1 mM EDTA, 1 mM DTT, 5 mM MgCl₂ and 500 μ M GTP or buffer alone. GTP hydrolysis was determined by measuring the release of phosphate at the times indicated²¹. Guanine nucleotide exchange rates for G α_{i3} were determined by measuring the time dependence of GTP γ S binding to the protein²⁰.

Received 29 July; accepted 15 August 1996.

- Dohlman, H., Apanieski, D., Chen, Y., Song, J. & Nusskern, D. *Mol. Cell Biol.* **15**, 3635–3643 (1995).
- Koelle, M. R. & Horvitz, H. R. *Cell* **84**, 115–125 (1996).
- Druey, K. M., Blumber, K. J., Kang, V. H. & Kehrl, J. H. *Nature* **379**, 742–746 (1996).
- Wong, Y. H. *et al. Nature* **351**, 63–65 (1991).
- Bogaski, M. S. & McCormick, F. *Nature* **366**, 643–645 (1993).
- Camps, M. *et al. Nature* **360**, 684–686 (1992).
- Katz, A., Wu, D. & Simon, M. I. *Nature* **360**, 686–689 (1992).
- Crespo, P., Xu, N., Simonds, W. F. & Gutkind, J. S. *Nature* **369**, 418–420 (1994).
- van-Biesen, T. *et al. Nature* **376**, 781–784 (1995).
- Kunkel, M. T. & Peralta, E. G. *Cell* **83**, 443–449 (1995).
- Arshavsky, V. & Bownds, M. D. *Nature* **357**, 416–417 (1992).
- Berstein, G. *et al. Cell* **70**, 411–418 (1992).
- Ausubel, F. *et al. in Current Protocols in Molecular Biology* (ed. Janssen, K.) 1314 (Wiley, New York, 1995).
- Gyuris, J., Golemis, E. A., Chertkov, H. & Brent, R. *Cell* **75**, 791–803 (1993).
- Gietz, D., St Jean, A., Woods, R. A. & Schiestl, R. H. *Nucleic Acids Res.* **20**, 1425 (1992).

- Kunkel, T., Roberts, J. & Zakour, R. *Methods Enzymol.* **154**, 367–382 (1987).
- Fields, T. A. & Casey, P. J. *J. Biol. Chem.* **270**, 23119–23125 (1995).
- Linder, M. E., Ewald, E., Miller, R. J. & Gilman, A. G. *J. Biol. Chem.* **265**, 8243–8251.
- Sternweis, P. C. & Robishaw, J. D. *J. Biol. Chem.* **259**, 13806–13813 (1984).
- Graziano, M. P., Freissmuth, M. & Gilman, A. G. *J. Biol. Chem.* **264**, 409–418 (1989).
- Higashijima, T., Ferguson, K. M., Smigel, M. D. & Gilman, A. G. *J. Biol. Chem.* **262**, 757–761 (1987).
- Hong, J. X., Wilson, G. L., Fox, C. H. & Kehrl, J. H. *J. Immunol.* **150**, 3895–3904 (1993).
- De-Vries, L., Mousli, M., Wurmser, A. & Farquhar, M. G. *Proc. Natl Acad. Sci. USA* **92**, 11916–11920 (1995).

ACKNOWLEDGEMENTS. T.W.H. and T.A.F. contributed equally to this work. We thank M. Linder for the G α_{i3} , C. Diesing for technical assistance, our colleagues for comments and A. Gilman for communication of unpublished results. T.W.H. was supported in part by a graduate student fellowship from Merck Research Laboratories. This work was supported by grants from the American Cancer Society to P.J.C. and the NSF Presidential young investigator program to E.G.P.

CORRESPONDENCE and requests for materials should be addressed to P.J.C. (e-mail: pjc@galactose.mc.duke.edu) or E.G.P. (e-mail: peralta@husc.harvard.edu).

Activation of the Raf-1 kinase cascade by coumermycin-induced dimerization

Michael A. Farrar*, José Alberola-Ila*†
& Roger M. Perlmutter*†‡

† Howard Hughes Medical Institute, and the Departments of * Immunology, ‡ Biochemistry and Medicine (Medical Genetics), University of Washington, Seattle, Washington 98195, USA

THE Raf-1 serine/threonine kinase is a key component of the MAP kinase cascade¹⁻³, regulating both proliferation and commitment to cell fate^{4,5}. Raf activation is stimulated following its translocation to the plasma membrane, a process that ordinarily requires interaction with the membrane-localized GTPase, Ras-GTP⁶⁻¹⁰. To investigate the mechanisms underlying Raf activation, we have developed a coumermycin-induced chemical dimerization method. We find that dimerization is by itself sufficient, in the absence of any membrane components, both to activate a modified Raf protein and to stimulate the MAP kinase cascade appropriately. As Ras-GTP-induced membrane localization increases the effective intracellular Raf concentration, our results indicate

that homotypic oligomerization may ordinarily act to promote Raf activation *in vivo*.

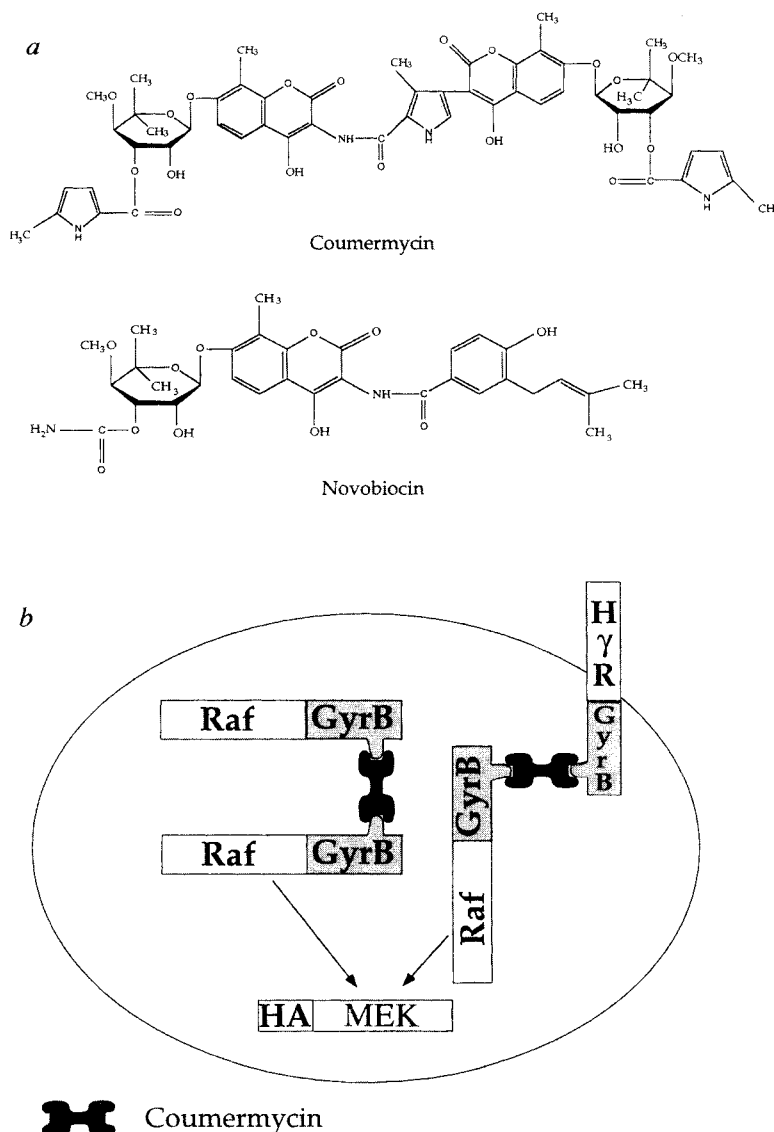
We used the *Streptomyces* product coumermycin because it has many of the characteristics of an ideal, exogenous dimerizer. Extensive structure-function studies have localized the binding target of coumermycin to the amino-terminal 24K subdomain of the B subunit of bacterial DNA gyrase¹¹ (GyrB). Coumermycin binds GyrB with a stoichiometry of 1:2, whereas a related 'monomeric' antibiotic, novobiocin (Fig. 1a), binds GyrB as a 1:1 complex¹². Thus coumermycin appears to act as a natural dimerizer of GyrB.

To generate a Raf kinase that could be inducibly targeted to the plasma membrane by coumermycin, we created chimaeric Raf fusion proteins in which the GyrB subunit was linked to the carboxy terminus of Raf (Fig. 1b). A membrane-anchored form of GyrB was generated by replacing the cytoplasmic domain of the human interferon- γ receptor (H γ R) with the GyrB subdomain (Fig. 1b). We transiently transfected simian COS cells with these expression constructs, together with a third construct encoding a haemagglutinin-tagged version of the Raf-1 target MAP-kinase kinase-1 (HA-MEK). Activation of the Raf-GyrB fusion protein was assessed after addition of coumermycin by measuring the *in vitro* kinase activity of immunoprecipitated HA-MEK.

Control experiments using COS cells transfected with either the combination of HA-MEK and wild-type Raf, or with HA-MEK and H γ R-GyrB demonstrated that coumermycin does not activate

FIG. 1 a, Chemical structures of coumermycin A1 and novobiocin. b, Experimental system, showing chimaeric fusion proteins of the GyrB polypeptide fragment fused to the C terminus of human c-Raf-1 or a cytoplasmically truncated human interferon- γ -receptor (H γ R), which were transiently expressed in COS cells with HA-MEK. Addition of coumermycin should lead to the formation of Raf-GyrB:H γ R-GyrB heterodimers, driving Raf-GyrB to the plasma membrane, and of Raf-GyrB homodimers, which should stay in the cytoplasm. Activation of Raf-GyrB was monitored by assaying HA-MEK activity.

METHODS. A cDNA encoding the N-terminal 24K polypeptide fragment of bacterial DNA gyrase²⁴ was obtained using the polymerase chain reaction (PCR) with DH5 α bacterial DNA as a template. A double-stranded oligonucleotide encoding a flexible 18-amino-acid peptide (GSEGGKSSGSGSESKVDTS)²⁵ was added to the 3' end of the GyrB cDNA. All constructs involving GyrB included this linker sequence. The modified GyrB cDNA was fused via an XbaI linker to the 3' end of a cDNA encoding human c-Raf-1. The H γ R-GyrB construct was prepared by inserting an XbaI linker just 3' of the H γ R transmembrane domain (following Lys 255) and then subcloning the GyrB cDNA fragment into this XbaI site. All constructs were then subcloned into the plasmid expression vectors pcDNA3 (Invitrogen) for use in transient transfections or pSR α , for generating stable transfectants. Sequences were verified using Sequenase II (US Biochemical) and standard techniques.



MEK or Raf directly (data not shown). In contrast, addition of coumermycin to cells triply transfected with plasmids encoding Raf-GyrB, H_YR-GyrB, and HA-MEK revealed a dose-dependent induction of HA-MEK activity. Maximal activation occurred at a concentration of 90 μ M coumermycin, yielding levels of Erk-2 substrate phosphorylation nearly identical to those obtained when cells were stimulated with phorbol myristate acetate (PMA), a potent activator of the MAP kinase signalling pathway (Fig. 2a). Although these initial results were consistent with the interpretation that Raf-GyrB was being activated by membrane localization, subsequent experiments demonstrated that cells transfected with only Raf-GyrB and HA-MEK remained coumermycin-responsive (Fig. 2b). These results suggested that dimerization, in the absence of membrane localization, activated Raf-GyrB.

To explore this possibility, we generated stable NIH3T3 cell lines expressing Raf-GyrB. Kinetic studies demonstrated that coumermycin stimulation leads to Raf-GyrB-induced MEK activation within 30 s (Fig. 3a), with more than 30-fold induction occurring 15 min after drug addition. The stimulatory effect was transient, decaying almost to baseline values within 4 hours. This desensitization was not due to inactivation of coumermycin on extended culture, because supernatants from cells stimulated for 4 h with coumermycin could still induce high levels of MEK activity when applied to unstimulated cells (data not shown). In general, the activation profile achieved with coumermycin closely paralleled that obtained when using PMA as stimulus (Fig. 3a).

We observed activation of Raf-GyrB in the presence of less than 10 nM coumermycin, with maximal responses at $\sim$ 1 μ M (Fig. 3b, c). This level of induction compared favourably with that obtained when cells were treated with either 20% fetal calf serum or PMA. In both transient and stable transfection systems, high doses of coumermycin inhibited MEK activation (Figs 2b, 3b and c), a result consistent with the hypothesis that coumermycin activates Raf-GyrB by dimerization, as at high coumermycin concentrations excess antibiotic should prevent dimer formation. Supporting this view, novobiocin, the 'monomeric' coumarin antibiotic, did not induce Raf-GyrB activation (Fig. 3d) and indeed it blocked coumermycin-induced Raf-GyrB activation when present in excess (Fig. 3d). Taken together, these results

indicate that activation of Raf-GyrB proceeds through antibiotic-induced dimerization.

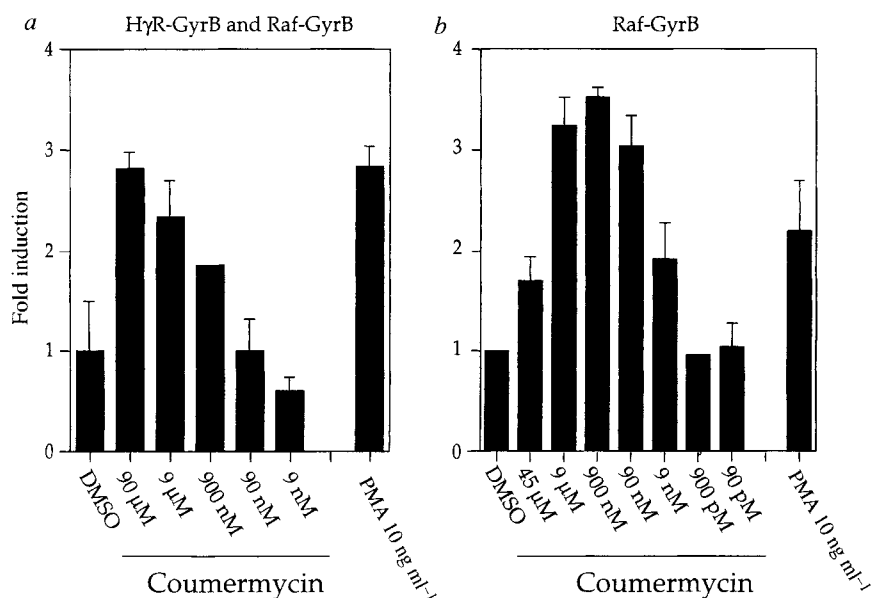
We next investigated the structural features of Raf required for coumermycin-induced activation of Raf-GyrB chimaeras by using transient transfection. First, a mutation that abrogates the catalytic activity of Raf (K375M, in which a methionine is substituted for a lysine at position 375)¹³, completely blocked coumermycin-induced Raf-GyrB activation (Fig. 4a). In contrast, a mutant that no longer interacts with Ras (R89L)¹⁴, and therefore should not be recruited to the plasma membrane, responded just as well to coumermycin as did wild-type Raf-GyrB (Fig. 4a). Furthermore, in these COS cells, replacement by phenylalanine of the two tyrosine residues (YY340/341FF) that are required for activation of Raf by *src*-family kinases¹⁵ had no effect on coumermycin-induced activation (Fig. 4a). Together these results suggest that activation of Raf-GyrB by coumermycin by-passes membrane localization entirely.

To confirm that membrane localization is not involved in coumermycin-induced Raf-GyrB activation, we generated membrane-free cytosolic extracts (S100 supernatants) from COS cells transiently transfected with Raf-GyrB and HA-MEK-encoding plasmids. Coumermycin stimulation of these extracts resulted in MEK activation equivalent to that in stimulated intact cells (compare Figs 2 and 4b). Furthermore, maximal stimulation in S100 supernatants required coumermycin concentrations tenfold less than those used for intact cells, probably reflecting more effective exposure of the chimaeric protein to the antibiotic. These results show that dimerization-driven activation of Raf-GyrB does not require either membrane localization or the addition of any membrane-localized cofactors.

Raf binds directly to Ras-GTP *in vivo* but this interaction does not stimulate Raf kinase function *in vitro*^{6,8}. Hence some other process at the membrane, perhaps involving tyrosine phosphorylation, has been proposed to activate Raf. Support for this hypothesis emerged when artificial tethering of Raf to the plasma membrane, achieved through carboxy-terminal farnesylation, was shown to promote constitutive activation of Raf¹⁶⁻¹⁸. *In vitro* activation of Raf has also been achieved by addition of membrane preparations to partially purified Raf, in a process

FIG. 2 Coumermycin induces activation of Raf-GyrB fusion proteins in COS cell transfectants. a, COS cells triply transfected with H_YR-GyrB, Raf-GyrB and HA-MEK were stimulated for 30 min with the indicated amounts of coumermycin or PMA (10 ng ml⁻¹) and analysed for activation of HA-MEK by *in vitro* kinase assay using catalytically inactive Erk-2 (Erk2_{K375R} containing an arginine for lysine replacement at position 52) as substrate²⁶. b, COS cells were transfected with Raf-GyrB and HA-MEK, stimulated with coumermycin or PMA, and then analysed for HA-MEK activation. Results shown are the means for duplicate determinations, normalized for HA-MEK abundance, and are representative of 3 experiments.

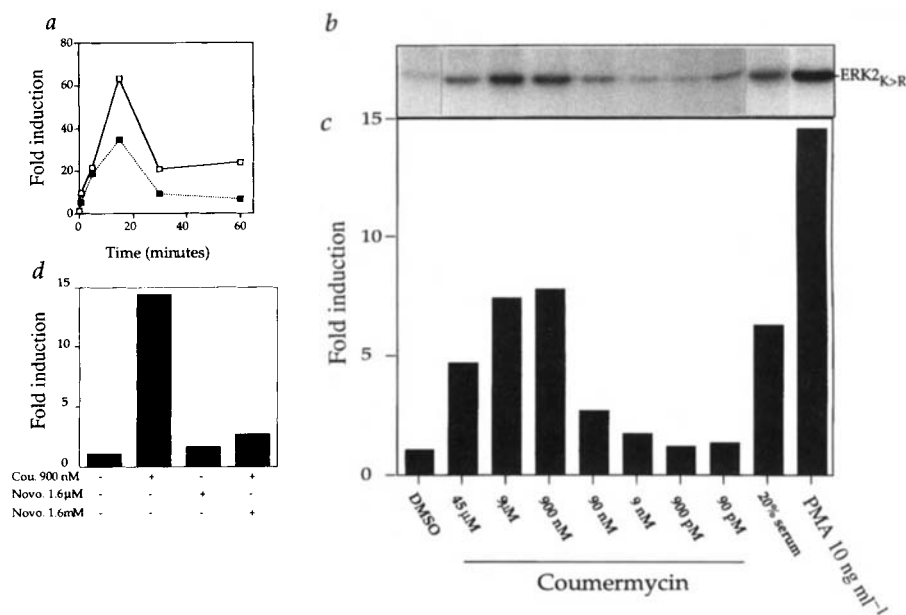
METHODS. a, b, COS cells were transfected in 100-mm dishes with 3 μ g ml⁻¹ of each plasmid using DEAE-dextran²⁷. After 48 h, cells were washed with PBS and placed in serum-free Dulbecco's modified Eagle's medium (DMEM) for a further 18 h, then stimulated with the indicated amounts of coumermycin (Sigma) or PMA (Sigma), dissolved in dimethylsulphoxide (DMSO), for 30 min. Mek-1 *in vitro* kinase assays were done as described²⁶. *In vitro* kinase reaction products were separated by 10% SDS-PAGE under reducing conditions and then transferred to nitrocellulose (Schleicher & Schuell). The amount of radioactivity incorporated into the Erk2_{K375R} substrate was quantified using a Molecular Dynamics phosphorimager and ImageQuant software (Molecular Dynamics). To normalize for the amount of HA-MEK precipitated and loaded onto the gel, blots were probed with polyclonal



rabbit anti-MEK antisera (Transduction Labs) in conjunction with horseradish-peroxidase-conjugated donkey anti-rabbit antiserum (Amersham) and developed by chemiluminescence (Amersham). Western blots were quantified using a Bio-Rad imaging densitometer and Molecular Analyst Software (Bio-Rad).

FIG. 3 Coumermycin induces a dose-dependent, transient activation of MEK in stable Raf-GyrB transfectants in NIH3T3 cells. **a**, Kinetics of coumermycin-induced MEK activation. Cells were stimulated with 900 nM coumermycin (■) or 10 ng ml⁻¹ PMA (□) and assayed at the indicated times. Results are means for duplicate determinations and representative of three experiments. **b**, Dose-response of coumermycin-induced Raf-GyrB activation. Cells were stimulated for 5 min with the indicated amounts of coumermycin, PMA or fetal bovine serum. Results were obtained using a single, representative clone. **c**, Quantification of MEK activity after normalization for total MEK protein in immunoprecipitations. Results are means for duplicate experiments, repeated 3 times. **d**, Blockade of coumermycin activation with novobiocin. NIH3T3 transfectants were stimulated for 5 min with either dimethylsulphoxide (DMSO) (lane 1), coumermycin (900 nM, lane 2), novobiocin (1.6 µM, lane 3) or novobiocin (1.6 mM) for 15 min, followed by addition of coumermycin (900 nM) for an additional 5 min (lane 4). Complete blockade of coumermycin-induced MEK activity was also obtained with 16 µM novobiocin.

METHODS. NIH3T3 cells were transfected using calcium phosphate²⁸ with the Raf-GyrB-encoding plasmid vector pSRα-Raf-GyrB. After selection in 1 mg ml⁻¹ G418, cells were cloned by limiting dilution. Clones were screened for expression by western blotting using an anti-GyrB monoclonal antibody (7D3) and two representative clones were chosen. For analysis,



cells were plated in 100-mm dishes, serum-starved for 18 h, then stimulated with varying doses of coumermycin, novobiocin, PMA (10 ng ml⁻¹), DMSO or 20% FBS (Gemini). MEK activity was assayed as for Fig. 2, except that an anti-MEK antibody (Transduction Labs) was used to immunoprecipitate cellular MEK-1.

that is ATP-dependent, enhanced by activated Src and Ras, and sensitive to resident phosphatases^{19,20}.

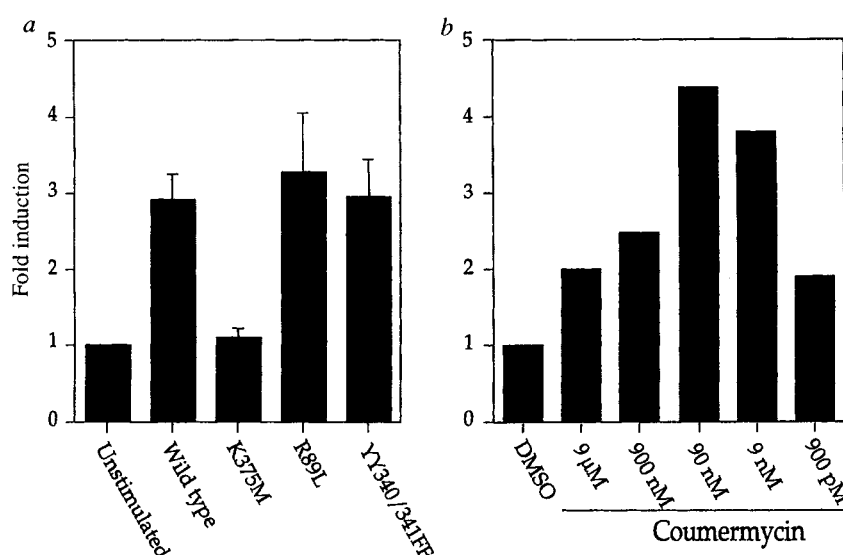
These observations indicate that Raf is activated at the plasma membrane, but they do not fully explain how. Although oligomerization has not yet been shown to stimulate activation of wild-type Raf²¹, our demonstration that Raf kinase activity can be induced by artificially dimerizing modified Raf proteins suggests that membrane localization of Raf, ordinarily achieved through binding to Ras-GTP, may promote clustering of the kinase, leading to activation. This localization of Raf could thus serve two purposes:

to render the kinase accessible to membrane kinases and phosphatases that regulate function, and to increase the local concentration of Raf by depositing it on a spatially restricted two-dimensional surface.

Our coumermycin/GyrB dimerization strategy may be applicable to intracellular signalling generally. Unlike the macrolide dimerizer FK1012 (ref. 22), coumermycin is commercially available, and its behaviour in cells and animals has been thoroughly investigated²³. The successful activation of Lck-GyrB chimaeras using coumermycin (S. Tucker, A. Norment, M.A.F. and R.M.P.,

FIG. 4 Coumermycin-dependent activation of Raf-GyrB does not require interaction with Ras, tyrosine phosphorylation or membrane localization. **a**, COS cells were transfected with Raf-GyrB or with plasmid constructs bearing defined mutations in Raf-GyrB. Transfected cells were stimulated for 30 min with 900 nM coumermycin or DMSO and HA-MEK activity assayed. Western blotting of cell lysates confirmed that equal amounts of mutant protein were present in both stimulated and unstimulated conditions. Bars represent mean $\pm$ s.e.m. for $n \geq 4$, except for K375M ($n = 3$). **b**, Raf-GyrB can be activated in the absence of membrane components. COS cells were transfected with HA-MEK and Raf-GyrB and lysed in assay buffer²⁶. S100 supernatants were prepared and then stimulated for 15 min at 30 °C in the presence of DMSO, or increasing amounts of coumermycin. HA-MEK was immunoprecipitated and MEK *in vitro* kinase activity assessed.

METHODS. **a**, Raf-GyrB mutants were prepared by overlap extension PCR, subcloned into pcDNA3, and then sequenced to verify that only the desired mutations had been introduced. For western blotting of Raf-GyrB mutants, either an anti-Raf antibody (Transduction Labs) was used or an antibody that specifically recognizes an epitope found within the 24K GyrB polypeptide (7D3). **b**, COS cells were transiently transfected with vectors encoding Raf-GyrB and HA-MEK. 32 h after transfection, cells were washed with PBS and placed in serum-free DMEM for further 18 h. After serum starvation, cells were trypsinized, washed twice with ice-cold PBS to remove trypsin, and lysed in assay



buffer containing 2 mM leupeptin, 400 mM PMSF and 20 µg ml⁻¹ aprotinin. Cell lysates were sonicated three times and then centrifuged at 100,000g for 20 min in a Beckman TLX ultracentrifuge. Supernatants (S100) were pooled and redistributed into 100-µl aliquots. Aliquots were stimulated for 15 min at 30 °C with DMSO or coumermycin, and HA-MEK activity was assayed after immunoprecipitation²⁶.

unpublished results), together with the Raf-GyrB results presented here, indicates that many kinases will prove tractable in this system. Furthermore, favourable pharmacological properties of coumermycin should enable the implementation of the dimerization-based signalling strategies in whole animal models. □

Received 28 March; accepted 5 July 1996.

- Howe, L. R. et al. *Cell* **71**, 335–342 (1992).
- Dent, P. et al. *Science* **257**, 1404–1407 (1992).
- Kyriakis, J. M. et al. *Nature* **358**, 417–421 (1992).
- Marshall, C. J. *Cell* **80**, 179–185 (1995).
- Dickson, B. & Hafen, E. *Curr. Opin. Genet. Dev.* **4**, 64–70 (1994).
- Zhang, X. et al. *Nature* **364**, 308–313 (1993).
- Wame, P. H., Vician, P. R. & Downward, J. *Nature* **364**, 352–355 (1993).
- Moodie, S. A., Willumsen, B. M., Weber, M. J. & Wolfman, A. *Science* **260**, 1658–1661 (1993).
- Van Aelst, L., Barr, M., Marcus, S., Polverino, A. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6213–6217 (1993).
- Vojtek, A. B., Hollenberg, S. M. & Cooper, J. A. *Cell* **74**, 205–214 (1993).
- Gilbert, E. J. & Maxwell, A. *Molec. Microbiol.* **12**, 365–373 (1994).
- Ali, J. A., Jackson, A. P., Howells, A. J. & Maxwell, A. *Biochemistry* **32**, 2717–2724 (1993).
- Morrison, D. K., Heidecker, G., Rapp, U. R. & Copeland, T. D. *J. biol. Chem.* **268**, 17309–17316 (1993).

- Fabian, J. R., Vojtek, A. B., Cooper, J. A. & Morrison, D. K. *Proc. natn. Acad. Sci. U.S.A.* **91**, 5982–5986 (1994).
- Fabian, J. R., Daar, I. O. & Morrison, D. K. *Molec. cell. Biol.* **13**, 7170–7179 (1993).
- Stokoe, D., Macdonald, S. G., Cadwallader, K., Symons, M. & Hancock, J. F. *Science* **264**, 1463–1467 (1994).
- Leever, S. J., Paterson, H. F. & Marshall, C. J. *Nature* **369**, 411–414 (1994).
- Marais, R., Light, Y., Paterson, H. F. & Marshall, C. J. *EMBO J.* **14**, 3136–3145 (1995).
- Dent, P., Reardon, D. B., Morrison, D. K. & Sturgill, T. W. *Molec. cell. Biol.* **15**, 4125–4135 (1995).
- Dent, P., Jelinek, T., Morrison, D. K., Weber, M. J. & Sturgill, T. W. *Science* **268**, 1902–1906 (1995).
- Wartmann, M. & Davis, R. J. *J. biol. Chem.* **269**, 6695–6701 (1994).
- Spencer, D. M., Wandless, T. J., Schreiber, S. L. & Crabtree, G. R. *Science* **262**, 1019–1024 (1993).
- Godfrey, J. C. & Price, K. E. *Adv. appl. Microbiol.* **15**, 653–718 (1972).
- Maxwell, A. *Molec. Microbiol.* **9**, 681–686 (1993).
- Pantoliano, M. W. et al. *Biochemistry* **30**, 10117–10125 (1991).
- Seeger, R. et al. *J. biol. Chem.* **269**, 25699–25709 (1994).
- Gonzalez, A. L. & Joly, E. *Trends Genet.* **11**, 216–217 (1995).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1989).

ACKNOWLEDGEMENTS. We thank J. Campbell for the Erk2_{kin} substrate, M. Gellert for the anti-GyrB monoclonal antibody, N. Ahn for the HA-MEK expression plasmid, D. Morrison for the human c-Raf-1 cDNA, J. Jackson for generating the initial GyrB cDNA, R. Peet for technical assistance, K. Prewitt for secretarial assistance and our colleagues for discussion. M.A.F. is supported by the Cancer Research Institute. R.M.P. is an Investigator of the Howard Hughes Medical Institute.

CORRESPONDENCE and requests for materials should be addressed to R.M.P. (e-mail: rmp@nucleus.immunol.washington.edu).

Oligomerization activates c-Raf-1 through a Ras-dependent mechanism

Zhijun Luo*, Guri Tzivion*, Peter J. Belshaw†, Demetrios Vavvas*, Mark Marshall‡ & Joseph Avruch*

* Diabetes Unit and Medical Services, Massachusetts General Hospital, and the Department of Medicine, Harvard Medical School, Boston, Massachusetts 02129, USA

† Howard Hughes Medical Institute, and Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA

‡ Department of Medicine, Division of Hematology and Oncology, Department of Biochemistry and Molecular Biology, and Walther Oncology Center, Indiana University, Indianapolis, Indiana 46202, USA

THE c-Raf-1 proto-oncoprotein is a Ras-GTP-regulated protein kinase¹ that associates *in situ* with 14-3-3 proteins^{2,3}, which are naturally dimeric^{4,5}. In COS cells, recombinant Raf is found in oligomeric assemblies. To examine whether induced oligomerization can alter Raf kinase activity, sequences encoding the FK506-binding protein FKBP12 were fused to the amino terminus of c-Raf-1, introducing a binding site for FK506. Oligomerization of recombinant FKBP-Raf *in situ*, induced by the addition of the dimeric FK506 derivative FK1012A, activated Raf kinase activity at least half as well as epidermal growth factor (EGF). As with EGF, activation of FKBP-Raf by FK1012A is entirely Ras-GTP dependent. Thus oligomerization of Raf *per se* promotes Raf activation through a Ras-dependent mechanism.

We have observed previously that N-terminally truncated, dimerization-deficient 14-3-3 ζ polypeptides, although able to bind c-Raf polypeptide *in situ* nearly as well as the full-length 14-3-3 ζ isoform, were recovered only in association with inactive Raf³. This suggested that the ability of 14-3-3 protein to dimerize might provide one important precondition for Raf activation, by enabling the apposition of inactive Raf bound to 14-3-3 ζ with a potential activator bound to the second 14-3-3 protein. To determine whether multimeric Raf complexes are present *in situ*, recombinant glutathione *S*-transferase (GST) or a GST-Raf fusion protein was expressed in COS cells together with a Myc-tagged Raf polypeptide. GSH-Sepharose purification of GST-

Raf from extracts of serum-deprived cells resulted in co-isolation of Myc-Raf, whereas no Myc-Raf was recovered with GST lacking Raf sequences (Fig. 1a). Treatment of cells with EGF activated both GST-Raf and Myc-Raf (data not shown) and gave an upshift in Raf polypeptide mobility on SDS-polyacrylamide gel electrophoresis (SDS-PAGE), but did not alter the amount of Myc-Raf that copurified with GST-Raf (Fig. 1a). Thus recombinant Raf exists in multimeric assemblies *in situ*, independently of EGF treatment.

To determine whether induced oligomerization of Raf could increase Raf activity, we constructed chimaeric FKBP12-Raf molecules (Fig. 1b), which can be oligomerized using the synthetic organic dimerizer FK1012A (Fig. 1b). The FKBP12 polypeptide is an abundant basic polypeptide of relative molecular mass 12,000 (*M_r* 12K) which contains a single binding site for the cell-permeant immunosuppressive drug FK506; the drug-protein complex binds to and inhibits calcineurin⁶. FK1012A is constructed by introduction of a crosslinker into the region of FK506 necessary for inhibition of calcineurin by the FKBP12/FK506 complex, a modification that simultaneously enables covalent dimerization of FK506 to FK1012A (Fig. 1b) to occur and eliminates its ability to interact with calcineurin but not FKBP^{7,8}.

FKBP-Raf chimaeric polypeptides, tagged with either a haemagglutinin (HA) or a Myc epitope, were expressed transiently in COS cells, separately or together; anti-HA immunoprecipitates contain FKBP-Myc-Raf (Fig. 1c) and anti-Myc immunoprecipitates contain FKBP-HA-Raf (Fig. 1d) only when the two variants are coexpressed. The FKBP-Raf oligomers found in serum-deprived cells require the Raf sequences (Fig. 1d). The amount of FKBP-Myc-Raf coprecipitating with FKBP-HA-Raf is unaffected by EGF (Fig. 1c), but is increased substantially in immunoprecipitates from cells pretreated with FK1012A (Fig. 1c, d). Thus FK1012 induces oligomerization of recombinant FKBP-Raf polypeptides *in situ*.

FKBP-HA-Raf has a low activity when assayed *in vitro* after extraction from serum-deprived COS cells, and is rapidly activated by treatment of the cells with EGF before extraction (Fig. 2a). Addition of FK1012A to serum-deprived cells also increases FKBP-Raf activity *in vitro*, but more slowly, rising to a steady-state level 30 min after addition of drug to intact cells (Fig. 2a). Optimal activation of FKBP-Raf is obtained at 0.3 μ M FK1012A (Fig. 2b), with modest inhibition at higher concentrations (data not shown); optimal activation of FKBP-Raf by FK1012A ranges from 40 to 100% of that achieved by maximal doses of EGF (Fig. 2). FK1012A has no effect on the basal activity of a Myc-epitope-tagged

Raf polypeptide (lacking FKBP12 sequences) (Fig. 2b), and gives modest inhibition of the EGF-induced activation of Myc-Raf (Fig. 2b). Thus activation of FKBP-Raf by FK1012A depends on the presence of FKBP12 sequences in the chimera. Similarly, activation of FKBP-Raf by FK1012A also requires the dimeric FK506 product, inasmuch as 0.6 μ M monomeric FK506M (Fig. 2a) has no effect on FKBP-Raf activity in serum-deprived cells, and at higher concentrations the monomer substantially inhibits activation by the dimer (although in part nonspecifically) (Fig. 2c; compare lanes 2 and 4 to lanes 5 and 7). These results indicate that oligomerization of Raf polypeptides is sufficient to activate Raf kinase activity (Fig. 2c).

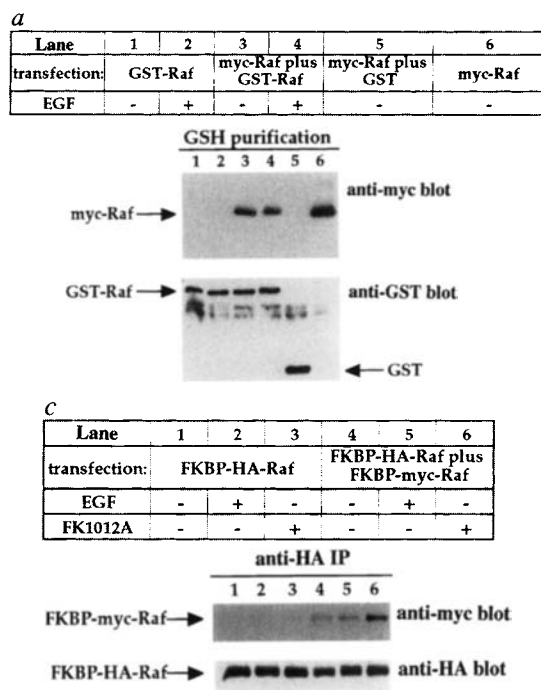
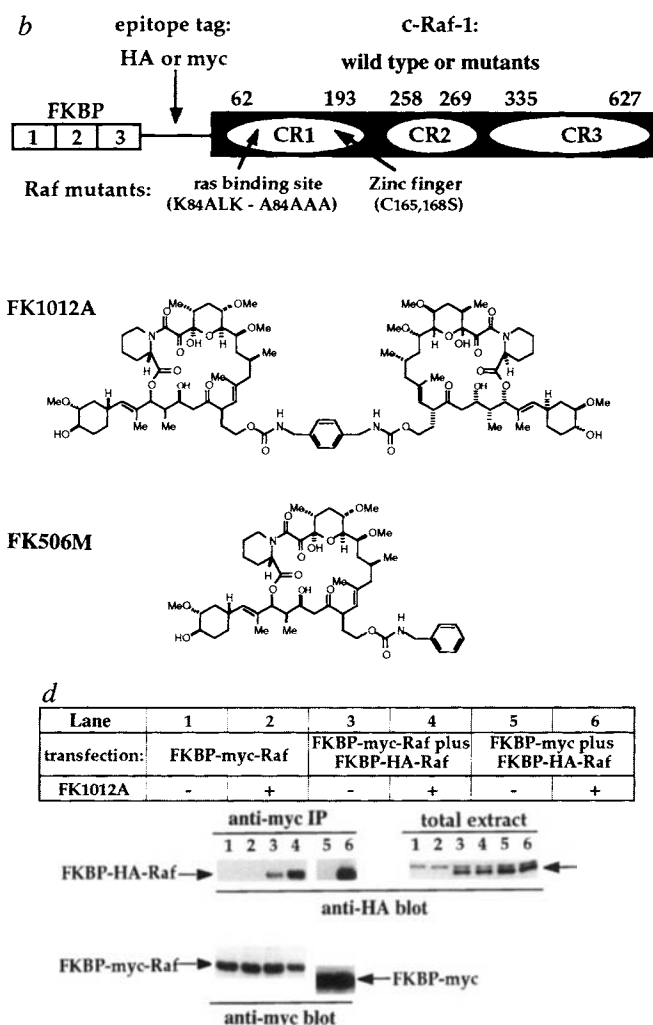


FIG. 1 Spontaneous and induced oligomerization of recombinant Raf polypeptides *in situ*. **a**, Recombinant Raf forms oligomers spontaneously in intact cells. Myc-Raf was with transiently expressed in COS cells, together with GST-Raf or GST alone. Purification of GST and GST-Raf on GSH-Sepharose beads (lanes 1–5) followed by Myc immunoblot shows Myc-Raf recovery only with GST-Raf (lanes 3, 4), but not with GST (lane 5); lane 6 contains a Myc-Raf standard. Treatment with EGF (50 ng ml⁻¹) for 30 min does not alter Raf oligomer abundance. **b**, Structure of the epitope-tagged FKBP-Raf polypeptides and the monomeric (FK506M) and dimeric (FK1012A) FK506 derivatives. 1, The chimaeric FKBP-epitope-c-Raf-1 protein contains 3 copies of the FKBP12 protein in series, linked to a protein epitope (Myc or HA) used for immunoprecipitation and immunoblotting by specific (anti-Myc or HA) monoclonal antibodies. The epitope in turn is fused to the N terminus of the human c-Raf-1 polypeptide; 2, the structure of monomeric FK506M, an FK506 derivatized to inhibit the binding of the FK506/FKBP complex to calcineurin, and the dimerizer FK1012A. **c**, The synthetic cell-permeant dimerizer FK1012A induces oligomerization of FKBP-Raf *in situ*. FKBP-HA-Raf was expressed transiently in COS cells alone or with FKBP-Myc-Raf. Serum-deprived cells were treated with carrier, EGF (50 ng ml⁻¹) or FK1012 (0.3 μ M) for 30 min before extraction. Anti-HA immunoprecipitates were immunoblotted with anti-HA (bottom) or anti-Myc (top) antibodies. **d**, Spontaneous, but not FK1012-induced, FKBP-Raf oligomers require Raf sequences. FKBP-HA-Raf was coexpressed in COS cells with FKBP-Myc-Raf (lanes 3, 4) or FKBP-Myc (lanes 5, 6). Serum-deprived cells were treated with FK1012 (0.3 μ M) for 30 min before extraction. Anti-Myc immunoprecipitates were immunoblotted with anti-HA antibody (top, left) and anti-Myc antibody (bottom, left). Cell extracts were immunoblotted with anti-HA antibody (top, right). **METHODS**. In **a**, transfections used 5 μ g DNA per plate, including 2.5 μ g of the vector pMT2 alone (lanes 1, 2) or encoding Myc-Raf (lanes 3–6) and 2.5 μ g of pEBG³ encoding GST-Raf (lanes 1–4) or GST (lane 5). The chimaeric FKBP-Raf polypeptides were all expressed in the vector pBJ5

We investigated whether the activation of FKBP-Raf by FK1012A was dependent on its interaction with Ras-GTP. The GTP charging of Ras can be partly inhibited by overexpression of N17-Ras⁹, with a consequent inhibition in the EGF-induced activation of FKBP-Raf. Notably, N17-Ras also results in a comparable inhibition of the FK1012A-induced activation of FKBP-Raf (Fig. 3a). An alternative, independent approach to interrupt the interaction between Ras and Raf is through mutation of either of the two domains on Raf that mediate the two-site interaction with Ras *in situ*^{10–13} (Z.L. *et al.*, manuscript submitted). Mutation of the Raf N-terminal Ras-binding domain (by conversion of Raf K84ALK to A84AAA) or the Raf zinc-finger (by



(ref. 18). The FKBP, called Fpk, is human FKBP12 containing the mutations G89I and P90K, which abrogate the ability of the FKBP-FK506 complex to bind to calcineurin¹⁷. The parent construct is MFpk3E^{7,20}, modified to remove the N-terminal Src myristoylation motif. Three copies of Fpk are fused, followed by an HA²¹ or Myc²² epitope, introduced as a *Sall*/*EcoRI* fragment using synthetic oligonucleotides. The human c-Raf-1 cDNA was introduced 3' to the epitope at the *EcoRI* site. The FKBP-Myc used in **d** lacks Raf sequences. The synthesis of FK506M and FK1012A have been described⁷. Transfection of COS M7 cells used the DEAE-dextran method. A total of 5 μ g of plasmid DNA was used. After 48 h, the cells were deprived of serum for an additional 18 h before the treatments indicated. Cells were extracted into Tris-Cl (50 mM, pH 7.5), NaCl (0.1 M), β -glycerophosphate (50 mM), EDTA (1 mM), EGTA (1 mM), DTT (1 mM), NaVO₄ (2 mM), Triton-X100 (1%) and proteinase inhibitors. Recombinant Raf polypeptides were immunoprecipitated from portions equalized for protein content, by incubation for 1 h at 4 °C with the monoclonal anti-HA antibody 12CA5 (ref. 21), or the anti-Myc antibody 9E10.2 (ref. 22). Immunoblots of the recombinant epitope-tagged polypeptides used the ECL method, with the antibodies indicated.

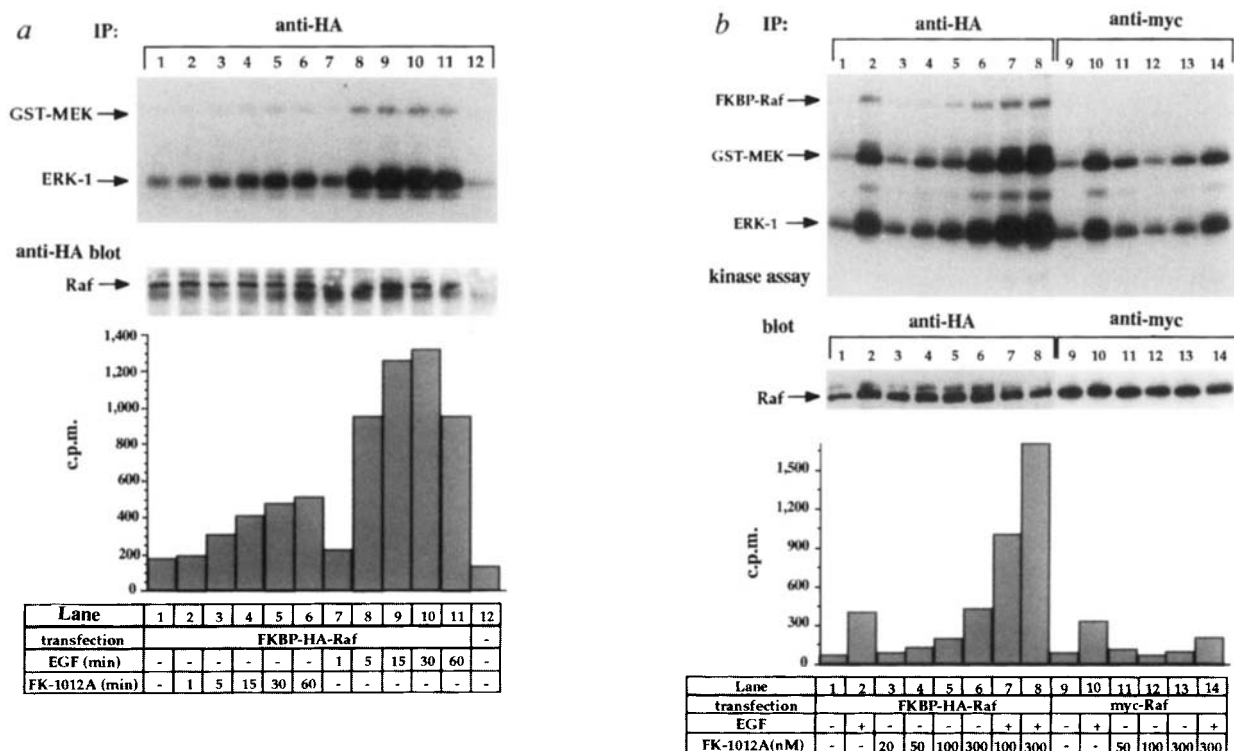
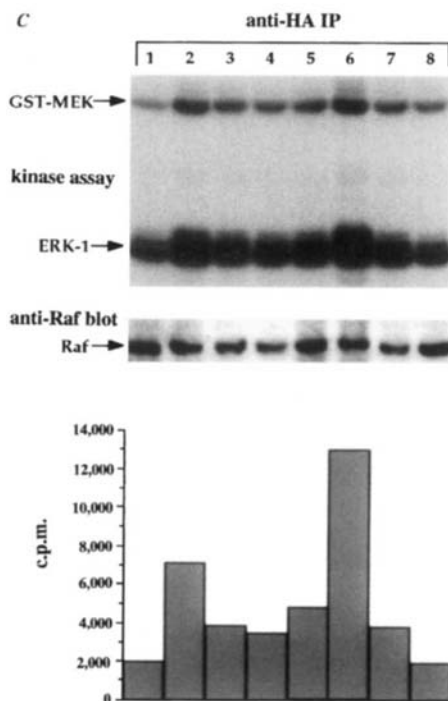


FIG. 2 FK1012A activates FKBP-Raf *in situ*. **a**, Time course of Raf activation after addition of FK1012A or EGF. COS cells transiently expressing FKBP-HA-Raf were deprived of serum 48 h after transfection; 18 h later, the cells were treated with carrier (lane 1), FK1012A (0.3 μ M, lanes 2–6) or EGF (50 ng ml⁻¹, lanes 7–11) for the times indicated. Cell extracts equalized for protein content were immunoprecipitated with the anti-HA monoclonal antibody 12CA5. The Raf kinase activity in the immune precipitates was measured in a two-stage assay involving the sequential transfer of ³²P from [γ -³²P]ATP to prokaryotic recombinant GST-MEK1, and ERK-1. The reaction mixtures were separated by SDS-PAGE, transferred to PVDF membranes, exposed for autoradiography (top), anti-HA immunoblot (middle) and liquid scintillation counting of the excised ERK-1 band (bottom); lane 12 shows the Raf kinase assay after a mock transfection. **b**, The concentration-dependent action of FK1012A on FKBP-HA-Raf and Myc-Raf compared with maximal EGF. Serum-deprived COS cells transiently expressing FKBP-HA-Raf (lanes 1–8) or Myc-Raf (lanes 9–14) were exposed to carrier (lanes 1, 9) FK1012A, at the concentrations indicated (lanes 3–6, 11–13). EGF (50 ng ml⁻¹) (lanes 2, 10) or both FK1012A and EGF together (lanes 7, 8, 14) for 30 min before extraction. Immunoprecipitation with anti-HA antibody 12CA5 (lanes 1–8) or anti-Myc antibody 9B7.3 was followed by assay for Raf kinase. Top, Raf kinase assay; middle, immunoblot of the immunoprecipitated Raf polypeptides using anti-HA antibody (12CA5) for lanes 1–8, and anti-Myc antibody 9E10.2 for lanes 9–14; bottom, ³²P incorporated into ERK-1 determined by liquid scintillation counting. **c**, Monomeric FK506M does not activate FKBP-HA-Raf, and inhibits activation by FK1012A. Serum-deprived COS cells transiently expressing FKBP-HA-Raf were treated with carrier (lane 1), FK1012A (0.3 μ M, lanes 2–4, 6) or EGF (50 ng ml⁻¹, lanes 5–7). FK506M monomer was added together with FK1012A (1 μ M, lane 3; or 3 μ M, lane 4) or EGF (3 μ M, lane 7), or alone (0.6 μ M, lane 8). Cells extracts were prepared 30 min later, and anti-HA immunoprecipitates were analysed for Raf kinase activity (top and bottom), and for FKBP-HA-Raf content by immunoblot with a polyclonal antipeptide antibody to the C-terminal 12 amino acids of human c-Raf-1.

METHODS. The washed immunoprecipitates were assayed for Raf kinase^{1,7,14} in a two-stage incubation at 30 °C in 0.1 ml, containing 0.1 mM [γ -³²P]-ATP (4,000 c.p.m. pmole⁻¹), 10 mM MgCl₂, and 0.2 μ g GST-MEK1. ERK-1 (1.0 μ g) was added 20 min after addition of the ATP; the reaction was terminated 30 min later by addition of SDS sample buffer. After SDS-PAGE, transfer to PVDF membranes and autoradiography, ³²P incorporation into ERK-1 was measured by liquid scintillation counting. Immunoblots were done as for Fig. 1, using the antibodies indicated.



Lane	1	2	3	4	5	6	7	8
EGF	-	-	-	-	+	+	+	-
FK-1012A (nM)	-	300	300	300	-	300	-	-
FK506-M (nM)	-	-	1,000	3,000	-	-	3,000	600

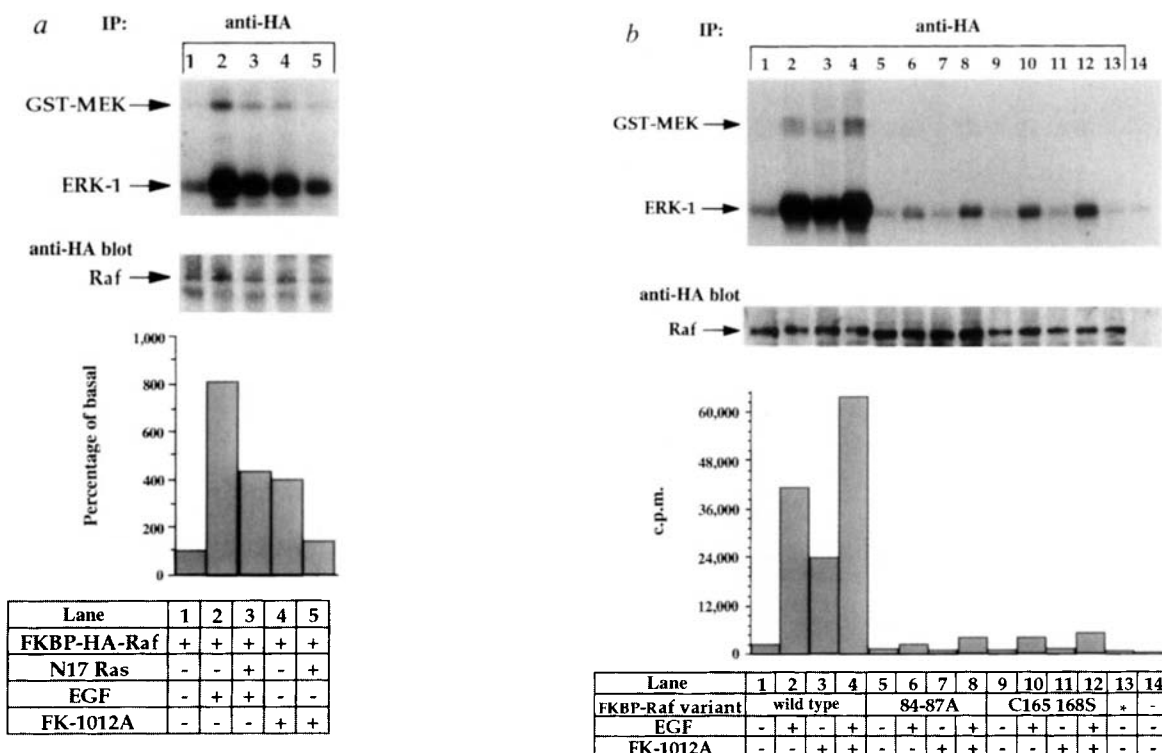


FIG. 3 Raf activation by the dimerizer, FK1012A, is Ras-GTP dependent. *a*, Asn-17 Ras inhibits FKBP-HA-RAF activation by EGF and FK1012A. Serum-deprived COS cells transiently expressing FKBP-HA-Raf alone (lanes 1, 2, 4), or together with Asn-17 Ha-Ras (lanes 3, 5), were treated with EGF (50 ng ml⁻¹; lanes 2, 3) or FK1012A (lanes 4, 5) for 30 min before extraction. Anti-HA immunoprecipitates were assayed for Raf kinase activity (top and bottom) and recombinant FKBP-HA-Raf content by immunoblot with anti-HA antibody (middle). *b*, Raf mutations that inhibit the binding of Raf to Ras *in situ* abolish the activation of FKBP-HA-Raf by EGF and FK1012A. Serum-deprived COS cells transiently expressing FKBP-HA-Raf with a wild-type Raf structure (lanes 1–4), FKBP-HA-Raf (K84 ALK to A84AAA, lanes 5–8), FKBP-HA-Raf (C165, 168S, lanes 9–12) FKBP-HA-Raf (K375M, lane 13) were treated with carrier (lanes 1, 5, 9, 13), EGF

(50 µg µl⁻¹; lanes 2, 6, 10), FK1012A (0.3 µM, lanes 3, 7, 11) or EGF plus FK1012A (lanes 4, 8, 12) for 30 min before extraction. Anti-HA immunoprecipitates were assayed for Raf kinase activity (top and bottom) as for Fig. 2, and immunoblotted for FKBP-HA-Raf content using anti-HA antibody. Lane 14 shows the autophosphorylation of the GST-MEK1 and ERK-1 substrates in the absence of Raf.

METHODS. In *a*, transfections contained 1 µg DNA encoding FKBP-HA-Raf, together with 5 µg of pCMV5-FLAG DNA, either as empty vector (lanes 1, 2, 4) or encoding Asn-17 Ha-Ras (lanes 2, 5). In *b*, transfections used 5 µg plasmid DNA encoding each of the FKBP-HA-Raf variants. Cell treatments, extraction, kinase assay and immunoblot were as in Figs 1 and 2, or as described.

conversion of Cys at 165 and 168 to Ser) each inhibits dramatically the ability of EGF to activate FKBP-Raf *in situ* without altering the levels of FKBP-Raf polypeptide expression after transient transfection. These mutations also abolish completely the ability of FK1012A to activate FKBP-Raf (Fig. 3*b*). Inasmuch as we have shown previously that these and related Raf mutations do not prevent activation of the Raf catalytic domain when an alternative, Ras-independent mechanism for Raf translocation to the plasma membrane is provided (Z.L. *et al.*, manuscript submitted), we conclude that activation of FKBP-Raf by FK1012A is entirely Ras-dependent. These results indicate that the ability of Ras-GTP to activate FKBP-Raf (and presumably endogenous Raf) is limited by the state of Raf oligomerization, and promotion of Raf oligomerization enhances the ability of Ras-GTP to activate Raf.

We therefore propose the following model for Raf activation. In serum-deprived cells, Raf exists in an oligomeric (probably dimeric) state assembled through the binding of Raf to the C terminus of 14-3-3 protein, which is itself dimeric. The 14-3-3 polypeptides are far more abundant than Raf, although it should be noted that they can bind many proteins apart from Raf. It is likely that only those Raf polypeptides that are in a dimer with another Raf polypeptide, or with another protein kinase that is capable of transphosphorylating Raf^{14,15}, will be susceptible to activation *in situ* by membrane-bound Ras-GTP. We suggest that the addition of FK1012A to cells expressing FKBP-Raf promotes FKBP-Raf activation primarily by oligomerizing FKBP-Raf/14-

3-3 complexes that each contain only one FKBP-Raf polypeptide. EGF does not regulate the abundance of Raf oligomers, but increases the GTP charging of Ras. Membrane-bound Ras-GTP serves as the activating ligand for the oligomeric Raf assembly, recruiting c-Raf-1 to the membrane and initiating a conformational change that enables the Raf polypeptide to interact productively with an activator, possibly a second Raf polypeptide or other kinase, perhaps through a transphosphorylation^{16,17}. This event may not in itself be sufficient for stable activation of c-Raf-1, so a further modification, perhaps contributed by a distinct, membrane-associated kinase or a transacylation, might be required to complete the activation of c-Raf-1. Conversely, in the case of B-Raf, it seems that adding prenylated Ras *in vitro* directly to the 14-3-3/B-Raf complex is sufficient to promote activation^{18,19}. Our results provide further insight into the complex regulation of the Raf kinases, and illustrate the ubiquitous role of oligomerization in signal transduction. □

Received 15 May; accepted 1 August 1996.

- Avruch, J., Zhang, X.-F. & Kyriakis, J. *Trends Biochem. Sci.* **19**, 279–283 (1994).
- Freed, E., Symons, M., McDonald, S. G., McCormick, F. & Ruggieri, R. *Science* **265**, 1713–1715 (1994).
- Luo, Z., Zhang, X.-F., Rapp, U. & Avruch, J. *J. Biol. Chem.* **270**, 23681–23687 (1995).
- Xiao, B. *et al.* *Nature* **376**, 188–191 (1995).
- Liu, D. *et al.* *Nature* **376**, 191–194 (1995).
- Liu, J. *et al.* *Cell* **68**, 807–815 (1991).
- Spencer, D. M., Wandless, T. J., Schreiber, S. L. & Crabtree, G. R. *Science* **262**, 1019–1024 (1993).
- Pruschy, M. N. *et al.* *Curr. Biol.* **1**, 163–172 (1994).
- Boguski, M. S. & McCormick, F. *Nature* **366**, 643–654 (1993).

10. Zhang, X.-F. et al. *Nature* **364**, 308–313 (1993).
11. Vojtek, A. B., Hollenberg, S. M. & Cooper, J. *Cell* **74**, 205–214 (1993).
12. Chuang, E. et al. *Mol. Cell. Biol.* **14**, 5318–5325 (1994).
13. Hu, C.-D. et al. *J. Biol. Chem.* **270**, 30274–30277 (1995).
14. Kovachina, K. S. et al. *J. Biol. Chem.* **265**, 12115–12118 (1990).
15. Dent, P., Jelinek, T., Morrison, D. K., Weber, M. J. & Sturgill, T. W. *Science* **268**, 1902–1906 (1995).
16. Lemmon, M. A. & Schlessinger, J. *Trends Biochem. Sci.* **19**, 459–463 (1994).
17. Johnson, L. N., Noble, M. E. M. & Owen, D. J. *Cell* **85**, 149–158 (1996).
18. Yamamoto, B. et al. *J. Biol. Chem.* **270**, 11723–11726 (1995).
19. Okada, T., Masuda, T., Shinkai, M., Kariya, K.-I. & Kataoka, T. *J. Biol. Chem.* **271**, 4671–4678 (1996).
20. Belshaw, P. J., Ho, S. N., Crabtree, G. R. & Schreiber, S. L. *Proc. Natl Acad. Sci. USA* **93**, 4604–4607 (1996).
21. Field, J. et al. *Mol. Cell. Biol.* **8**, 2159–2166 (1988).
22. Evan, G. I., Lewis, G. K., Ramsey, G. & Bishop, J. M. *Mol. Cell. Biol.* **5**, 3610–3616 (1985).

ACKNOWLEDGEMENTS. Z.L. and G.T. contributed equally to this work. We thank S. L. Schreiber for discussions and review of the manuscript, and J. Prendable for help in preparation of the manuscript. This work was supported by grants from E. Lilly Inc. (to M.M. and J.A.) and the American Cancer Society (to M.M.). P.J.B. was supported by a grant to S.L. Schreiber. Z.L. was supported by an NRSA from the NIH.

CORRESPONDENCE and requests for materials should be addressed to Z.L. (e-mail: Luo @helix.mgh.harvard.edu).

Transcription activation in cells lacking TAF_{II}s

Scott S. Walker, Joseph C. Reese, Lynne M. Apone & Michael R. Green

Howard Hughes Medical Institute, Program in Molecular Medicine, University of Massachusetts Medical Center, 373 Plantation Street, Suite 309, Worcester, Massachusetts 01605, USA

THE general transcription factor TFIID is composed of the TATA-box-binding protein (TBP) and a set of TBP-associated factors (TAF_{II}s) (ref. 1). *In vitro*, TAF_{II}s are required for activated transcription, and have been proposed to be obligatory targets of transcriptional activator proteins (activators)². The function of TAF_{II}s has not been investigated systematically *in vivo*. A *Saccharomyces cerevisiae* TAF_{II} complex (yTAF_{II} complex) has been identified that shares functional and structural similarities with higher eukaryotic TFIID. In particular, most yTAF_{II}s are the homologue of a higher eukaryotic TAF_{II} (refs 3, 4). Here we report that inactivation or depletion of six different yTAF_{II}s, including the core yTAF_{II} that contacts TBP, does not compromise transcriptional activation. We conclude that *in vivo*, activated transcription of many genes can occur in the absence of functional yTAF_{II}s, and that in these instances another transcription component(s) must be the target of the activator.

The TBP-associated factor yTAF_{II}145 is the homologue of higher eukaryotic TAF_{II}250, and is the only yTAF_{II} known directly to contact TBP³. For these and other reasons, yTAF_{II}145/yTAF_{II}250 is thought to be the core subunit of the complex^{1,5}. We isolated several temperature-sensitive (ts) alleles of yTAF_{II}145, which is encoded by an essential gene^{3,4}. Cells containing these yTAF_{II}145 mutant alleles did not grow at 37 °C, and displayed a rapid cessation of cell division when shifted from 23 °C to the non-permissive temperature (Fig. 1a). Immunoblot analysis (Fig. 1b) shows that, when shifted to 37 °C, the level of yTAF_{II}145 rapidly decreased, and by 4 h was undetectable in both temperature-sensitive strains. The levels of several other yTAF_{II}s, and to a lesser extent yTBP, also decreased. In contrast, components that are not part of the yTAF_{II} complex, such as Sua7 (yTFIIB), were unaffected by the temperature shift. These data suggest that loss of yTAF_{II}145 disrupts the yTAF_{II} complex, resulting in the degradation of other yTAF_{II}s.

To investigate the consequences of yTAF_{II} inactivation on transcription, we initially examined strains bearing temperature-

sensitive mutations in yTAF_{II}145 or TSM1, which is encoded by an essential gene and is the homologue of the higher eukaryotic TAF_{II}150 (ref. 6), or deleted of yTAF_{II}30. The gene encoding yTAF_{II}30, *ANCI*, is not essential, but cells lacking yTAF_{II}30 grow slowly and are inviable at 37 °C (ref. 7).

Yeast strains containing these TAF_{II} mutants or, as a control, a temperature-sensitive general transcription factor (GTF) mutant (TBP or RNA polymerase II), were shifted to 37 °C and RNA isolated at various times. Transcription was measured by S1 nuclease protection using probes specific for five endogenous yeast genes driven by diverse activators^{8–12}. For all five genes, as shown in Fig. 2a, in the *tbp*^{ts} (lanes 17–20) and *rpb*^{ts} (lanes 21–24) strains, shifting to the non-permissive temperature substantially reduced transcription, which by 4 h was undetectable. Similar results were obtained with an *srb4*^{ts} mutant strain (ref. 13, and data not shown). In contrast, even after 4 h at the non-permissive temperature, transcription was not significantly affected in the wild-type strain (lanes 1–4), strains bearing temperature-sensitive mutations in yTAF_{II}145 (lanes 5–12), TSM1 (lanes 13–16), or the strain that had yTAF_{II}30 deleted (lanes 25 and 26).

The genes examined in Fig. 2a were actively transcribed at the time of the temperature shift, leaving open the possibility that yTAF_{II}s might be required for the initiation, but not maintenance, of a transcriptionally active state. To examine this possibility, we tested whether yTAF_{II} inactivation affected the inducible transcription of the yeast copper metallothionein gene, *CUP1*. Tran-

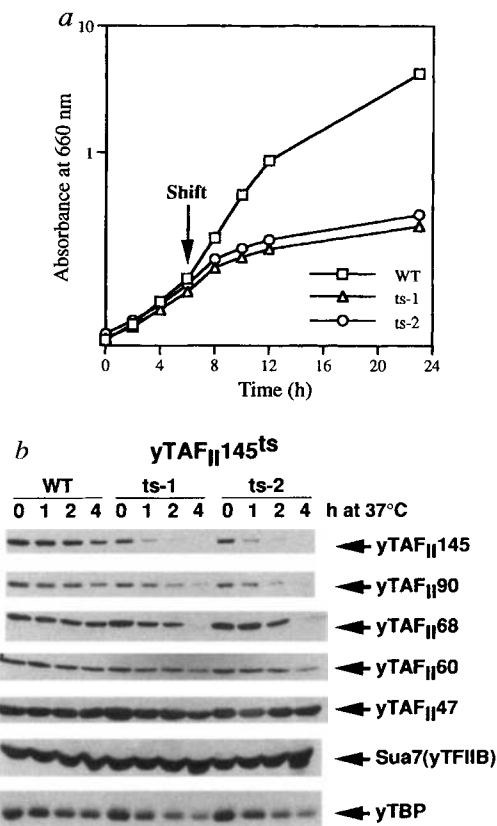


FIG. 1 Temperature-sensitive yTAF_{II}145 mutants. **a**, Rapid cessation of cell growth follows shifting *ytaf*_{II}145^{ts} cultures from 23 °C to 37 °C. Cells were grown in YPD and absorbance at 660 nm (*A*₆₆₀) was determined for each culture at various times before and after the temperature shift (arrow). **b**, Immunoblot of extracts from cells before and after the temperature shift. Portions of each culture were removed for small-scale whole-cell extract preparation, before (0) and 1, 2 and 4 h after the shift to 37 °C. Immunoblots were probed with the polyclonal antisera for the yTAF_{II} or GTF indicated on the right.

scriptional induction of *CUP1* is extremely rapid and involves the direct binding of Cu^{2+} to the acidic activator ACE1, converting it to an active form^{14,15}. Cultures of cells containing temperature-sensitive mutations in yTAF_{II}145, TSM1, TBP or RNA polymerase II were shifted to the non-permissive temperature for 4 h, exposed to copper sulphate (1 mM) for 30 min, and transcription was quantified (Fig. 2b, bottom). At the permissive temperature (23 °C), activation of *CUP1* was similar in all strains (Fig. 2b). At the non-permissive temperature (37 °C), induction of *CUP1* was severely diminished in the *tbp*^{ts} cells (lanes 19 and 20), and was undetectable in *rpb1*^{ts} cells (lanes 23 and 24). However, cells containing the yTAF_{II} mutations (*ytaf*_{II}145^{ts} and *tsm1*^{ts}) supported transcriptional induction of the *CUP1* gene to a level indistinguishable from that of wild type, even after incubation for 4 h at the non-permissive temperature (lanes 7, 8, 11, 12, 15 and 16). *CUP1* transcription was also normal in the *ytaf*_{II}30Δ strain (lanes 25–28).

To confirm and extend the results with the temperature-sensitive yTAF_{II} mutants, we depleted individually four yTAF_{II}s (yTAF_{II}145, yTAF_{II}90, yTAF_{II}68 and yTAF_{II}47) by using another experimental strategy. Of these, yTAF_{II}145 and yTAF_{II}90 have been previously described³, but yTAF_{II}47 and yTAF_{II}68 are new yTAF_{II}s, cloned on the basis of microsequence analysis of the corresponding subunits of the purified yTAF_{II} complex³. yTAF_{II}68 is homologous to higher eukaryotic hTAF_{II}20 and dTAF_{II}30α (data not shown), whereas yTAF_{II}47 has no recognizable higher eukaryotic homologue. The genes encoding both yTAF_{II}68 and yTAF_{II}47 are essential for viability (data not shown).

The gene for each yTAF_{II} was placed under regulation of the glucose-repressible yeast GAL1 promoter and introduced into a strain in which the corresponding endogenous yTAF_{II} gene had been deleted. We used these strains to analyse transcription after yTAF_{II} depletion (Fig. 3a and b; see below). After incubation in glucose-containing medium, each GAL1-regulatable yTAF_{II} was rapidly depleted, and by 12 h was undetectable (Fig. 3c–f). In several instances, and in particular for yTAF_{II}68, the directed depletion of a single TAF_{II} resulted in loss of other yTAF_{II} subunits, consistent with yTAF_{II}s existing *in vivo* as a multisubunit complex. Thus transcription was measured in cells lacking multiple yTAF_{II}s. From immunoblotting we estimate that yeast cells normally contain 3,000–6,000 copies of yTAF_{II}s, which was decreased more than 95% by glucose depletion (data not shown). The maximal 300 copies of each yTAF_{II} remaining is far below the ~4,000 actively transcribed genes in a yeast cell¹⁶.

Transcriptional induction of *CUP1* was examined in these conditional strains 12 h after glucose-mediated yTAF_{II} shutoff (Fig. 3a, b, bottom). In cells in which yTAF_{II}145, yTAF_{II}90, yTAF_{II}68 or yTAF_{II}47 had been deleted, *CUP1* transcription was identical to that of the wild-type strain (Fig. 3a).

Heat shock is another well characterized, inducible transcription response¹⁷. Strains containing glucose-regulatable alleles of yTAF_{II}145, yTAF_{II}90, yTAF_{II}68, or yTAF_{II}47 were examined for transcriptional induction of the *SSA4* gene¹⁸. For each strain, yTAF_{II} depletion did not diminish transcriptional activation (Fig. 3b).

Although temperature-sensitive inactivation of yTAF_{II}145 or TSM1 did not result in a general defect in transcription activation (Fig. 2a, b), a rapid growth arrest was evident (Fig. 1a; ref. 19 and data not shown). The ability rapidly to inactivate these temperature-sensitive yTAF_{II}s enabled

us to gain insight into the nature of the growth arrest. When the *ytaf*_{II}145^{ts} strain was shifted to 37 °C, there was an accumulation of large, unbudded cells (Fig. 4a), indicative of a G1-phase cell-cycle arrest²⁰. However, at the non-permissive temperature, the *tsm1*^{ts} strain arrested as budded cells, characteristic of the G2/M phase (Fig. 4a). To delineate the cell-cycle block more precisely, the arrested cells were subjected to fluorescence-activated cell sorting (FACS). *ytaf*_{II}145^{ts} cells arrested with a 1N DNA content (Fig. 4b), consistent with a block in G1, and analogous to results with the mammalian TAF_{II}250 homologue²¹; conversely, *tsm1*^{ts} cells had a 2N DNA content, consistent with a block in G2.

Taken together, our results indicate that activated transcription can occur in the absence of multiple yTAF_{II}s. This conclusion is based on two independent strategies functionally to inactivate yTAF_{II}s: temperature-sensitive mutations, and conditional depletion. Moreover, a similar conclusion has been obtained using a third yTAF_{II} inactivation strategy (K. Struhl, personal communication). Particularly compelling are the results with yTAF_{II}145. It is the only yTAF_{II} known to contact TBP directly³, and its higher eukaryotic homologue, TAF_{II}250, is always required to reconstitute TFIID activity *in vitro*⁵. Therefore, for those genes tested, TAF_{II}s are not the obligatory targets of activators, and so in these instances another transcription component (s), such as TFIIB or TBP²², must serve this function.

Although inactivation of yTAF_{II}s did not compromise transcription of those genes examined, inactivation of different yTAF_{II}s led to distinct cell-cycle phenotypes. In yeast, cell-cycle progression requires the temporally regulated transcription of particular genes²³. An intriguing possibility is that yTAF_{II}s are not required

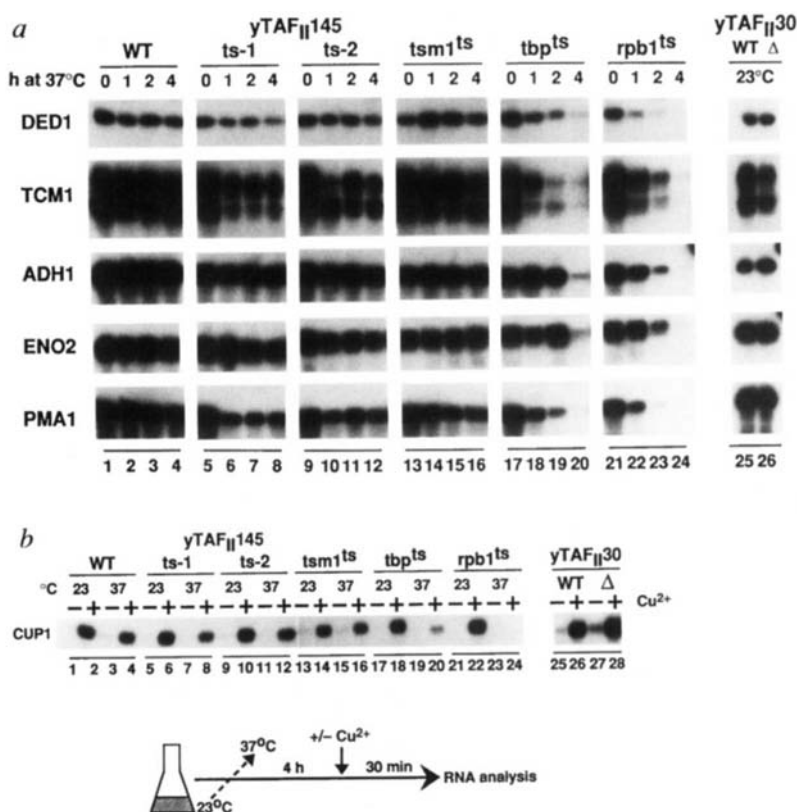
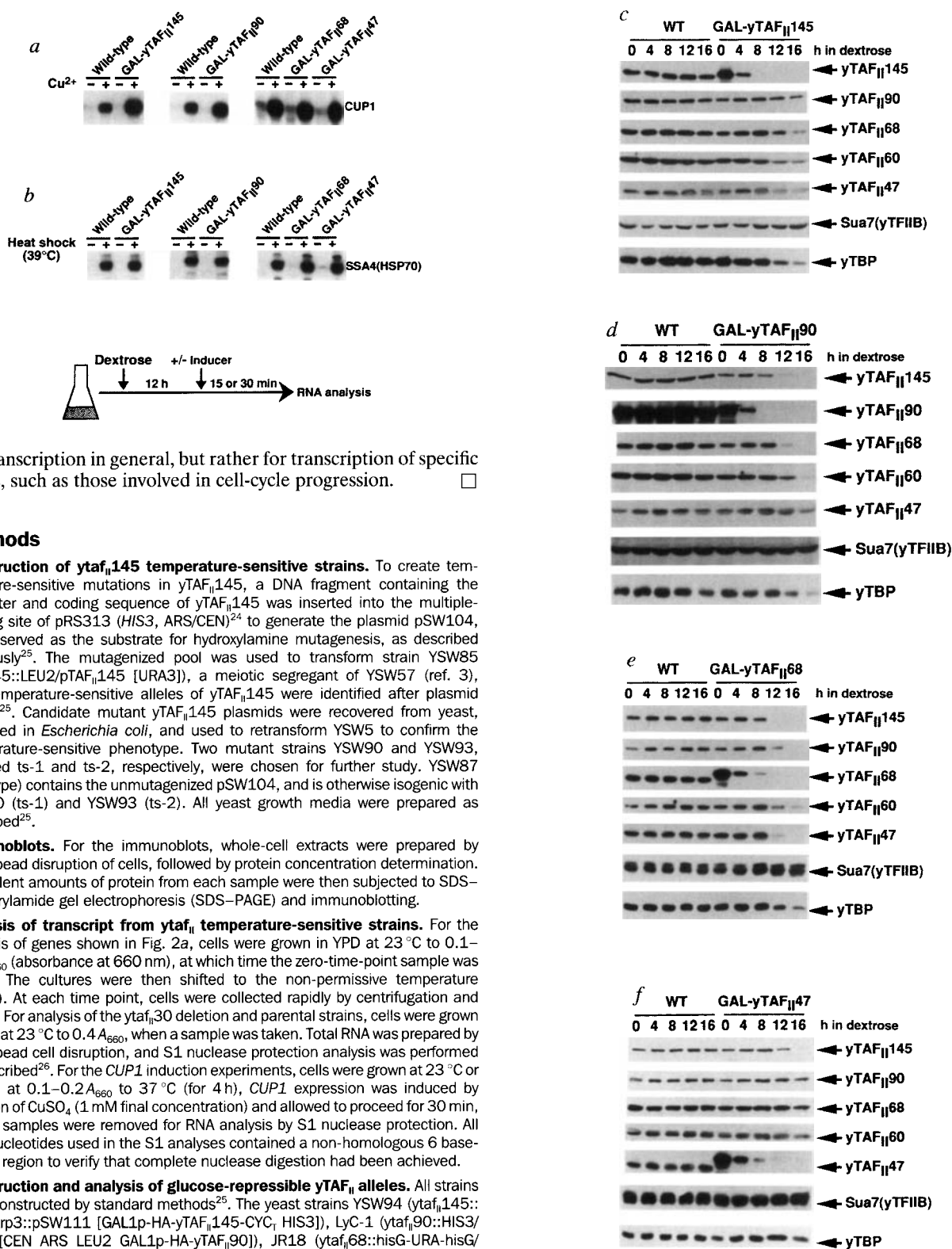


FIG. 2 Unabated transcription after yTAF_{II} inactivation. **a**, Transcription of several endogenous, chromosomal genes. For the *ytaf*_{II}145^{ts}, *tbp*^{ts} and *rpb1*^{ts} strains, culture samples were taken for RNA preparation and analysis before (0) and 1, 2 and 4 h after the shift to 37 °C. The *ytaf*_{II}30 deletion (Δ) and parental, wild-type (WT) cells were grown at 23 °C before RNA was prepared for analysis. The transcripts detected by S1 nuclease protection in each strain are indicated on the left. **b**, Transcriptional activation of *CUP1* in yTAF_{II} mutant strains. The experimental scheme is shown below the autoradiogram. Growth and induction of *CUP1* transcription in the *ytaf*_{II}30 deletion and wild-type strains was performed at 23 °C.



for transcription in general, but rather for transcription of specific genes, such as those involved in cell-cycle progression. □

Methods

Construction of *ytaf_{II}145* temperature-sensitive strains. To create temperature-sensitive mutations in *yTAF_{II}145*, a DNA fragment containing the promoter and coding sequence of *yTAF_{II}145* was inserted into the multiple-cloning site of pRS313 (*HIS3*, *ARS/CEN*)²⁴ to generate the plasmid pSW104, which served as the substrate for hydroxylamine mutagenesis, as described previously²⁵. The mutagenized pool was used to transform strain YSW85 (*taf_{II}145::LEU2/pTAF_{II}145 [URA3]*), a meiotic segregant of YSW57 (ref. 3), and temperature-sensitive alleles of *yTAF_{II}145* were identified after plasmid shuffle²⁵. Candidate mutant *yTAF_{II}145* plasmids were recovered from yeast, amplified in *Escherichia coli*, and used to retransform YSW5 to confirm the temperature-sensitive phenotype. Two mutant strains YSW90 and YSW93, denoted ts-1 and ts-2, respectively, were chosen for further study. YSW87 (wild type) contains the unmutagenized pSW104, and is otherwise isogenic with YSW90 (ts-1) and YSW93 (ts-2). All yeast growth media were prepared as described²⁵.

Immunoblots. For the immunoblots, whole-cell extracts were prepared by glass-bead disruption of cells, followed by protein concentration determination. Equivalent amounts of protein from each sample were then subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting.

Analysis of transcript from *ytaf_{II}* temperature-sensitive strains. For the analysis of genes shown in Fig. 2a, cells were grown in YPD at 23 °C to 0.1–0.2 *A*₆₆₀ (absorbance at 660 nm), at which time the zero-time-point sample was taken. The cultures were then shifted to the non-permissive temperature (37 °C). At each time point, cells were collected rapidly by centrifugation and frozen. For analysis of the *ytaf_{II}30* deletion and parental strains, cells were grown in YPD at 23 °C to 0.4 *A*₆₆₀, when a sample was taken. Total RNA was prepared by glass-bead cell disruption, and S1 nuclease protection analysis was performed as described²⁶. For the *CUP1* induction experiments, cells were grown at 23 °C or shifted at 0.1–0.2 *A*₆₆₀ to 37 °C (for 4 h), *CUP1* expression was induced by addition of CuSO₄ (1 mM final concentration) and allowed to proceed for 30 min, before samples were removed for RNA analysis by S1 nuclease protection. All oligonucleotides used in the S1 analyses contained a non-homologous 6 base-pair 3' region to verify that complete nuclease digestion had been achieved.

Construction and analysis of glucose-repressible *yTAF_{II}* alleles. All strains were constructed by standard methods²⁵. The yeast strains YSW94 (*ytaf_{II}145::LEU2/trp3::pSW111 [GAL1p-HA-yTAF_{II}145-CYC_I HIS3]*), LYC-1 (*ytaf_{II}90::HIS3/Lp16 [CEN ARS LEU2 GAL1p-HA-yTAF_{II}90]*), JR18 (*ytaf_{II}68::hisG-URA-hisG/trp3::pJE20 [GAL1p-HA-yTAF_{II}68-CYC_I HIS3]*) and JR11 (*ytaf_{II}47::hisG-URA3-hisG/trp3::pJR15 [GAL1p-HA-yTAF_{II}47-CYC_I HIS3]*) were used for conditional expression of *yTAF_{II}145*, *yTAF_{II}90*, *yTAF_{II}68*, and *yTAF_{II}47*, respectively. All GAL-*TAF_{II}* strains were incapable of growth in glucose-containing media (YPDextrose), confirming that these strains are solely dependent on the GAL-*yTAF_{II}* expression construct for viability. To deplete a *yTAF_{II}*, cells were first grown to 0.05 *A*₆₆₀ in YPGalactose (zero time point) and then shifted to YPDextrose. Portions were removed at the indicated times for immunoblot and RNA analysis. The *CUP1* transcript was detected by S1 nuclease protection and the *SSA4/HSP70* transcript was detected by primer extension. Induction of *CUP1* (by addition of CuSO₄ for 30 min) or *SSA4/HSP70* (by shifting the culture to 39 °C for 15 min) was performed 12 h after the media change.

Microscopy and FACS. For microscopic examination, cells were grown at 23 °C to 0.2 *A*₆₆₀, at which time a sample was removed, the cultures were shifted

FIG. 3 Activated transcription in cells with *yTAF_{II}*s depleted. Inducible transcription of a, *CUP1* and b, *SSA4* (*HSP70*) in YSW94, LYC1, JR18, and JR11. Immunoblot detection of: c, *yTAF_{II}145*; d, *yTAF_{II}90*; e, *yTAF_{II}68*; and f, *yTAF_{II}47*; in strains (YSW94, LYC1, JR18 and JR11, respectively), after incubation in YPDextrose. All four conditionally expressed *yTAF_{II}*s contain a haemagglutinin (HA)-epitope tag at their N terminus and thus have a slightly lower mobility than the native protein. For comparison in the immunoblots and transcription inductions, the 'wild-type' for YSW94 was strain YSW87 (*ytaf_{II}145::LEU2/pyTAF_{II}145 [ARS/CEN]*), for LYC1 it was LY3 (*ytaf_{II}90::HIS3/Lp6 [2 μ URA3 ADH2-yTAF_{II}90]*) and for strains JR18 and JR11 it was W303a.

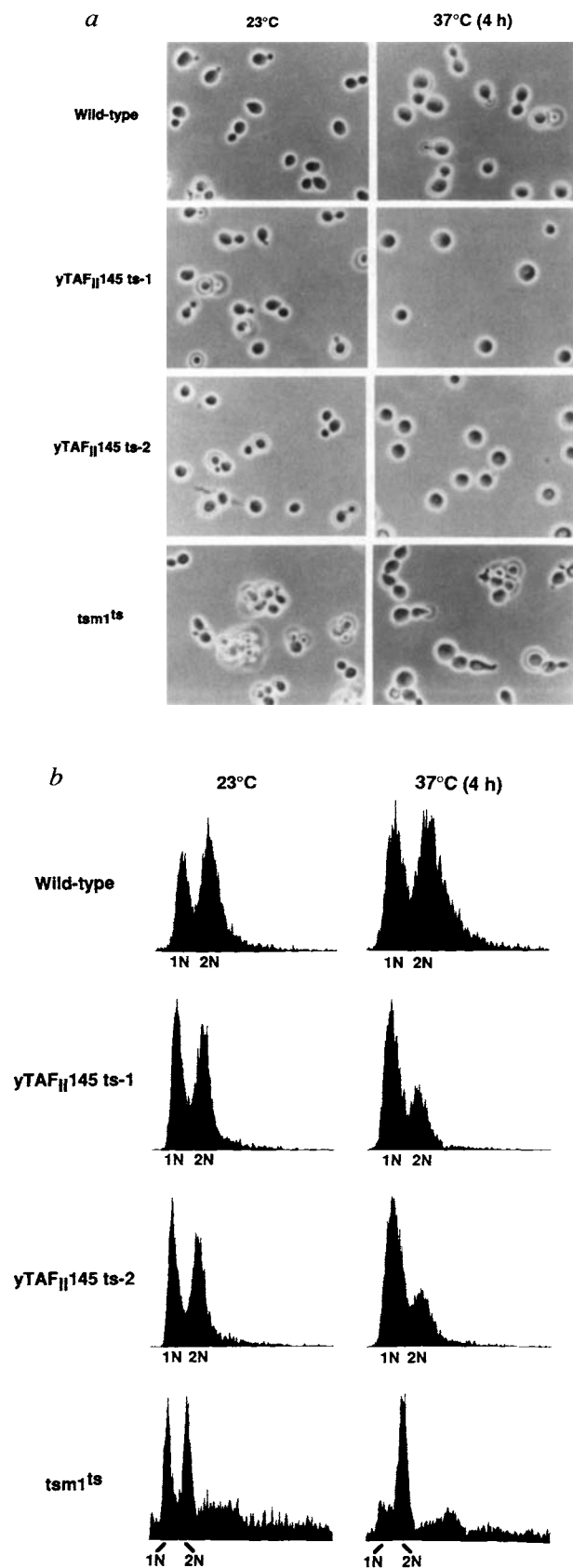


FIG. 4 Inactivation of yTAF_{II}145 and TSM1 leads to distinct cell-cycle phenotypes. *a*, Phase-contrast micrographs of wild type (YSW87), the two yTAF_{II}145 temperature-sensitive mutants, and the tsm1^{ts} mutant at 23 °C, and after 4 h at the non-permissive temperature (37 °C). All micrographs were taken at a magnification of $\times 1,000$. *b*, Flow cytometric (FACS) analysis, by propidium iodide staining, of the DNA content of wild-type, yTAF_{II}145^{ts} and tsm1^{ts} strains at 23 °C and after 4 h at 37 °C.

to 37 °C, and a sample of each was taken after 4 h. Cell samples were fixed in formaldehyde and mounted on poly-L-lysine-coated slides. For FACS analysis, cells were grown at 23 °C to 0.1–0.2 A_{600} , a sample was taken, and the culture was shifted to 37 °C for 4 h. Cells were prepared for FACS quantification of DNA content as previously described²⁷.

Received 9 May; accepted 16 July 1996.

1. Burley, S. D. & Roeder, R. B. *Annu. Rev. Biochem.* **65**, 769–799 (1996).
2. Goodrich, J. A. & Tjian, R. *Curr. Opin. Cell Biol.* **6**, 403–409 (1994).
3. Reese, J. C., Apone, L., Walker, S. S., Griffin, L. A. & Green, M. R. *Nature* **371**, 523–527 (1994).
4. Poon, D. *et al. Proc. Natl Acad. Sci. USA* **92**, 8224–8228 (1995).
5. Chen, J.-L., Attardi, L. D., Verrijzer, C. P., Yokomori, K. & Tjian, R. *Cell* **79**, 93–105 (1994).
6. Verrijzer, C. P., Yokomori, K., Chen, J.-L. & Tjian, R. *Science* **264**, 933–941 (1994).
7. Welch, M. D., Vinh, D. B. N., Okamura, H. H. & Drubin, D. G. *Genetics* **135**, 265–274 (1993).
8. Buchman, A. R. & Kornberg, R. D. *Mol. Cell. Biol.* **10**, 887–897 (1990).
9. Hamill, K. G., Nam, H. G. & Fried, H. M. *Mol. Cell. Biol.* **8**, 4328–4341 (1988).
10. Johnston, M. & Carlson, M. *The Molecular and Cellular Biology of the Yeast Saccharomyces* Vol. 2, Ch. 5 (eds Jones, W. W. *et al.*) (Cold Spring Harbor Laboratory Press, NY, 1992).
11. Kuo, M.-H. & Grayhack, E. *Mol. Cell. Biol.* **14**, 348–359 (1994).
12. Capieaux, E., Vignais, M. L., Sentenac, A. & Goffeau, A. *J. Biol. Chem.* **264**, 7437–7446 (1989).
13. Thompson, C. M. & Young, R. A. *Proc. Natl Acad. Sci. USA* **92**, 4587–4590 (1995).
14. Butt, T. R. *et al. Proc. Natl Acad. Sci. USA* **81**, 3332–3336 (1984).
15. Fürst, P., Hu, S., Hackett, R. & Hamer, D. *Cell* **55**, 705–717 (1988).
16. Lewin, B. *Genes* Vol. 4 (Cell, Cambridge, MA, 1990).
17. Bienz, M. & Pelham, H. R. B. *Adv. Genet.* **24**, 31–72 (1987).
18. Boorstein, W. R. & Craig, E. A. *J. Biol. Chem.* **256**, 18912–18921 (1990).
19. Mortimer, R. K. & Schild, D. *Microbiol. Rev.* **44**, 519–571 (1980).
20. Herskowitz, I. *Microbiol. Rev.* **52**, 536–553 (1988).
21. Nishimoto, T. *et al. Somat. Cell Genet.* **8**, 811–824 (1982).
22. Zawal, L. & Reinberg, D. *Annu. Rev. Biochem.* **64**, 533–561 (1995).
23. Koch, C. & Nasmyth, K. *Curr. Opin. Cell Biol.* **6**, 451–459 (1994).
24. Sikorski, R. S. & Hieter, P. *Genetics* **122**, 19–27 (1989).
25. Guthrie, C. & Fink, G. R. (eds) *Guide to Yeast Genetics and Molecular Biology* (Academic, New York, 1991).
26. Cormack, B. P. & Struhl, K. *Cell* **69**, 685–696 (1992).
27. Lew, D. J., Marinin, M. J. & Reed, S. I. *Cell* **69**, 317–327 (1992).

ACKNOWLEDGEMENTS. We thank J. Haber for yeast strain tsm1^{ts} (RY570); D. Drubin for the yTAF_{II}30 deletion strain DDY555 and its parent DDY547; K. Struhl for the yeast strain tpb^{ts} BY2 t.s.1; R. Young for the yeast strain rbp1^{ts} Y262; C. Peterson for yeast strains; K. Struhl, R. Young and C. Peterson for advice and communication of results before publication; D. Reinberg for antibodies; E. Craig for advice on the yeast heat-shock response; S. Dosey for the use of his microscope; M. Woda and the UMMC Flow Cytology Laboratory for FACS analysis; J. Kelly for technical assistance; and members of the Green laboratory for advice and encouragement. This work was supported by postdoctoral fellowships from the NIH and the American Cancer Society (Massachusetts Division) to S.S.W.; the Damon Runyon–Walter Winchell Cancer Research Foundation Fellowship (J.C.R.); and a grant from the NIH to M.R.G. M.R.G. is an investigator of the Howard Hughes Medical Institute.

CORRESPONDENCE and requests for materials should be addressed to M.R.G. (e-mail: michael.green@ummed.edu). Oligonucleotide sequences used are available from the authors on request. The sequences of yTAF_{II}47 (EMBL accession no. Z48483, ORF no. YP8132.02c) and yTAF_{II}68 (EMBL accession no. Z50046, ORF no. YD8358.02) are available with the completion of the yeast genome sequencing project.

TBP-associated factors are not generally required for transcriptional activation in yeast

Zarmik Moqtaderi*, Yu Bai†, David Poon†, P. Anthony Weil† & Kevin Struhl*

* Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

† Department of Molecular Physiology and Biophysics, Vanderbilt University Medical Center, Nashville, Tennessee 37232-0615, USA

THE transcription factor TFIID, a central component of the eukaryotic RNA polymerase II (Pol II) transcription apparatus, comprises the TATA-binding protein (TBP) and approximately ten TBP-associated factors (TAFs)¹. Although the essential role of TBP in all eukaryotic transcription has been extensively analysed *in vivo* and *in vitro*^{2,3}, the function of the TAFs is less clear. *In vitro*, TAFs are dispensable for basal transcription but are required for the response to activators¹. In addition, specific TAFs may act as molecular bridges between particular activators and the general transcription machinery^{4,5}. *In vivo*, TAFs are

required for yeast^{6,7} and mammalian⁸ cell growth, but little is known about their specific transcriptional functions. Using conditional alleles created by a new double-shutoff method, we show here that TAF depletion in yeast cells can reduce transcription from some promoters lacking conventional TATA elements. However, TAF depletion has surprisingly little effect on transcriptional enhancement by several activators, indicating that TAFs are not generally required for transcriptional activation in yeast.

We used a two-pronged approach to create strains with conditional TAF alleles in which the addition of copper ion leads to the simultaneous cessation of TAF messenger RNA synthesis and destruction of any TAF protein present in the cell (Fig. 1a). We generated strains with conditional alleles of TAF130(TAF145), TAF90, TAF60 and TAF19(Fun81), which are homologous to human TAF250, *Drosophila* TAF80, *Drosophila* TAF60 and human TAF18, respectively. Of these, TAF130 is particularly interesting, as it appears to be the scaffold on which the remaining TAFs assemble into the TFIID complex⁴. As controls, we generated strains containing conditional alleles of TFIIB and TBP. In all cases, the conditional knockout strains fail to grow on copper-containing medium (Fig. 1b), and the addition of copper ion causes cells to stop growing within about six hours (Fig. 1c). TAF90-depleted cells arrest frequently as large, budded cells, whereas cells depleted of other TAFs display variable and abnormal morphologies.

In general, transcription was analysed 8 hours after copper ion addition, when more than 95% of the cells are dead (they are unable to grow when returned to medium lacking copper). At this time, western blotting reveals that levels of TAF130, TAF90 and TBP are less than 5% of wild type (Fig. 1d). Although TAFs may not be completely eliminated by this procedure, they are reduced to less than 100–200 molecules per cell (data not shown; ref. 9), which is considerably less than the number of Pol II promoters per cell (~6,000).

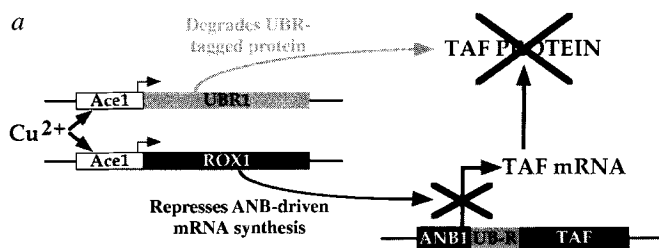


FIG. 1 Creation and characterization of strains with conditional TAF alleles. **a**, Strain construction. In the parent strain, ROX1²⁸, a transcriptional repressor, and UBR1²⁹, the N-end recognition protein involved in the ubiquitin degradation pathway, are expressed from a tightly regulated copper-inducible promoter²⁶. The TAF gene of interest is replaced by a derivative that contains an N-end recognition signal for rapid ubiquitin-dependent degradation and is driven by the ANB1 promoter, a target of Rox1. In the absence of copper ion, neither the repressor nor the protein degradation system is active, and the TAF is expressed. Upon copper addition, both the repressor and the ubiquitin degradation system are rapidly activated, leading to the simultaneous cessation of TAF mRNA synthesis and destruction of TAF protein. **b**, Growth of 10⁴ cells from each strain on synthetic complete medium either lacking (left) or containing (right) 500 μ M copper sulphate for two days at 30 °C. **c**, Growth curves of strains on synthetic complete medium to which 500 μ M copper sulphate was added at time zero. **d**, Western blot analysis of TAF depletion. Protein (30 μ g) from parental, TAF, and TBP depletion strains at various times after addition of 500 μ M CuSO₄ were probed with the relevant antibodies; dilutions of protein from the parental strain are indicated for TAF130 and TBP. Although full-length TAF130 (star) is very rapidly degraded, a truncated product (lacking a fragment of *M_r* ~ 25K), which may or may not have functional TAF activity, is degraded more slowly so that ~5% remains 8 h after copper addition.

When cells are grown under standard conditions, TAF depletion affects transcription of selected Pol II promoters (Fig. 2). Depletion of TAF130, TAF60, TAF90 and TAF19 does not significantly affect transcription of *ded1* or *his3 + 13*, promoters with canonical TATA elements^{10,11}. However, depletion of TAF130 significantly reduces the level of the *trp3* and *his3 + 1* transcripts, which arise from promoters with suboptimal, non-consensus TATA elements^{11,12}. This preferential effect on transcription from promoters containing weak TATA elements is also observed when TAF19 is depleted, albeit to a lesser extent and with slower kinetics, but it does not occur upon depletion of TAF90 or TAF60. Interestingly, the transcriptional pattern resulting from TAF130 or TAF19 depletion is similar to that mediated in yeast by human TBP, which has been suggested to interact inefficiently with yeast TAFs¹³. As expected, depletion of TBP or TFIIB results in a rapid and large reduction of all mRNA species tested. At a late time point (11 hours), depletion of TAF90 confers a moderate decrease of all transcripts. It is unclear whether this effect reflects a specific function of TAF90 or arises indirectly from cell death.

Surprisingly, when the conditional knockout strains are grown under conditions that support activation by Gcn4 or Ace1, TAF depletion does not significantly affect the level of activated transcription (Fig. 3a, b; Ace1-dependent activation appears slightly reduced upon TAF90 depletion). In contrast, depletion of TFIIB causes the loss of activated transcription in both situations. The observed Gcn4- and Ace1-activated transcription reflects initiation events that occur under conditions of TAF depletion, because mRNA half-lives are very short in comparison to the timescale of the experiment.

One explanation for the maintenance of Gcn4 and Ace1 activation after TAF depletion is that TAFs present in active transcription complexes might be preferentially sequestered from Ubr1-dependent degradation. To examine whether active transcription complexes could be assembled after TAF depletion, the

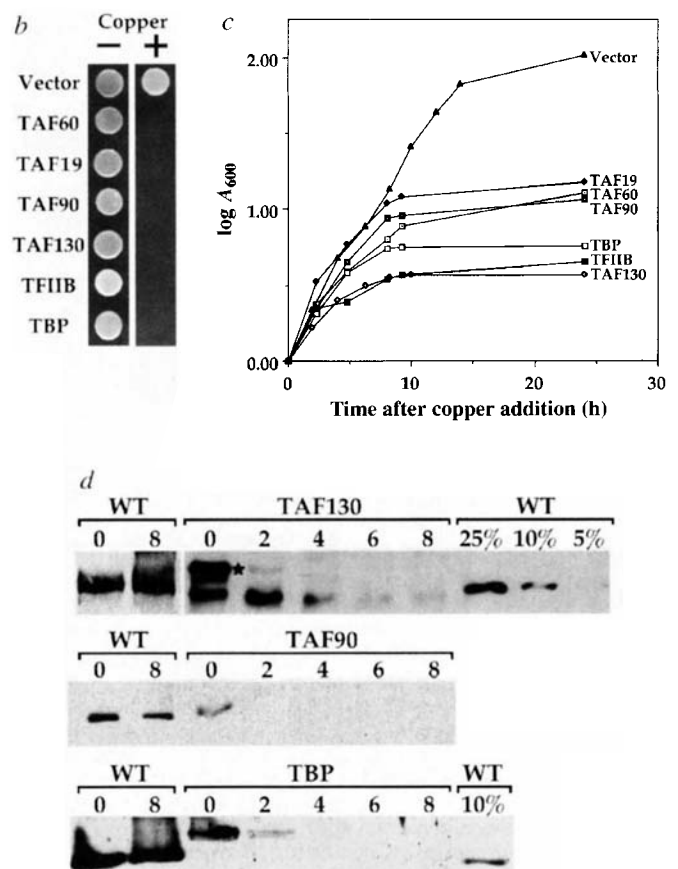
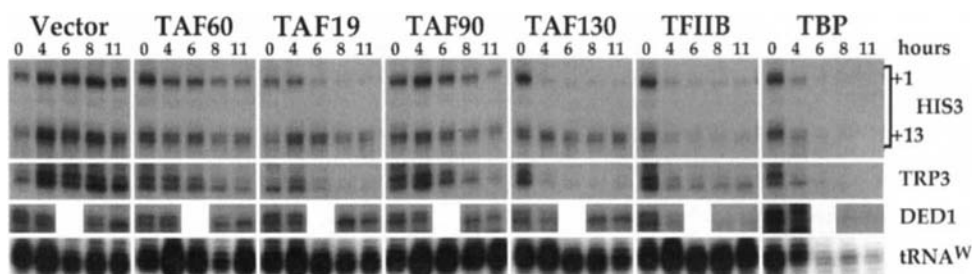


FIG. 2 Differential effects of TAF depletion on Pol II transcripts. Parental and TAF-depletion strains at the indicated time (hours) after the addition of copper were analysed for *his3*, *trp3*, *ded1* and *tRNA^W* transcription.



conditional knockout strains were grown in non-inducing conditions, treated with copper for 8 hours, and then tested for the ability to mediate activator-dependent transcription *de novo* (Fig. 4a, b). All the TAF-depleted strains show significant activation by Gal4 and heat-shock factor upon exposure to the relevant inducer, whereas activation is not observed in the TFIIB-depleted strain. Similarly, we observe efficient Ace1-dependent activation after TAFs were depleted either by placing the TAF genes under the control of the *GAL1,10* promoter and shifting cells to glucose (data not shown) or by shifting a *tsm1* (*dTAF150* homologue) strain to the restrictive temperature (Fig. 4c). The heat-shock and Ace1 activation responses in the TAF-depleted strains are comparable to the parental strain; Gal4 activation is reduced 3–4 fold. However, as very small fluctuations in Gal4 levels or changes in growth potential can have pronounced effects on Gal4-dependent transcription¹⁴, it is unclear whether the decrease reflects a mild activation defect or whether Gal4 expression is slightly perturbed for other reasons. Previously described activation-defective yeast strains are considerably more impaired for Gal4-dependent activation, and they are defective in the response to other acidic activators^{15–17}.

From the observation that TAF depletion does not significantly affect activation by Gcn4, Ace1, Gal4, Hsf, and unidentified activators involved in *ded1* and *his3* + 13 transcription, we conclude that the TAFs are not generally required for transcriptional

activation in yeast cells. This conclusion was reached independently in experiments where TAF depletion was obtained using temperature-sensitive mutants or a glucose shutoff procedure⁹. It is particularly striking that this is true of TAF130, which provides the scaffold for TAF assembly and without which TFIID is likely to be disrupted. Although TAFs are not generally required for transcriptional activation, they are essential for cell growth. One possibility is that TAFs are required for the response to a subset of activators that affect one or more essential genes. Alternatively, TAFs could subtly affect activation of many genes, such that the cumulative effects lead to cell inviability. Finally, as suggested by the effects on *trp3* and *his3* + 1 transcription, TAFs may be important for transcription from promoters with weak TATA elements.

Our conclusion is in apparent contrast to numerous experiments *in vitro*, which indicate that TAFs are crucial in all activated transcription. We do not favour the hypothesis that yeast TAFs

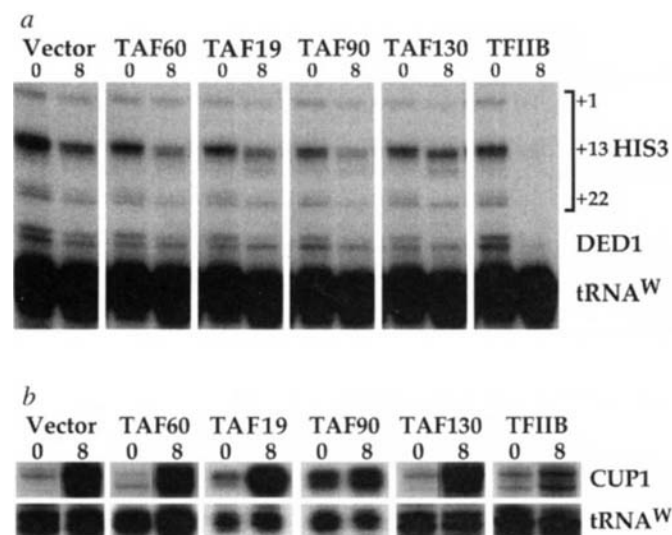


FIG. 3 Maintenance of Gcn4- and Ace1-activated transcription after TAF depletion. *a*, Parental and TAF-depletion strains cultured in glucose synthetic complete medium containing 10 mM 3-aminotriazole (a condition that induces Gcn4 synthesis) at 0 and 8 h after copper addition were analysed for *his3* and *ded1* transcription. Similar results were obtained using strains containing a plasmid that constitutively expresses Gcn4. *b*, Ace1 activation. Parental and TAF-depletion strains at 0 and 8 h after copper addition were analysed for *cup1* and *tRNA^W* transcription. Ace1 activation of *cup1* occurs simultaneously with the induction of TAF depletion.

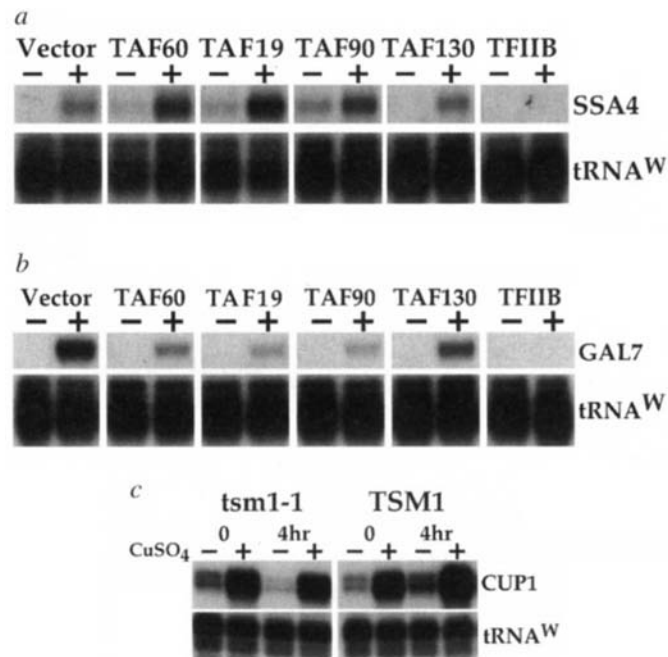


FIG. 4 Activation can be induced after TAF depletion. *a*, Activation by heat-shock factor. Parental and TAF-depletion strains at 8 h after copper addition that were (+ lanes) or were not (- lanes) subjected to a 12-min heat shock at 39.5 °C were analysed for *ssa4* and *tRNA^W* transcription. *b*, Activation by Gal4. Parental and TAF-depletion strains cultured in raffinose synthetic complete medium at 8 h after copper addition, which were or were not induced for 30 min with 2% galactose, were analysed for *gal7* and *tRNA^W* transcription. *c*, Activation by Ace1 after thermal inactivation of TSM1. Strain AW41³⁰ and a derivative containing a plasmid-borne TSM1 allele at 0 and 4 h after temperature shift which were or were not induced for 30 min with 500 μM copper sulphate were analysed for *cup1* and *tRNA^W* transcription. Growth of the *tsm1-1* strain ceases within 4 h after temperature shift.

are less important than their mammalian and *Drosophila* counterparts because: (1) TAFs are strongly conserved among eukaryotes; (2) TAF-dependent activation *in vitro* can be achieved with yeast components^{6,7}; and (3) activation can occur in a hamster cell line in which TAF250 (yeast TAF130 homologue) has been thermally inactivated (Fos transcription occurs normally, and it is unclear whether the reduction of cyclin A transcription is an indirect effect of cell-cycle arrest or a direct effect of TAF250)⁸. A more likely explanation is that TAFs are functionally redundant with other factors that are absent in typical *in vitro* reactions. Indeed, activated transcription in the apparent absence of TAFs can occur *in vitro* when reactions either contain Pol II holoenzyme^{18,19} or are performed on chromatin templates²⁰. Moreover, most *in vitro* transcription reactions are reconstituted with core Pol II, and hence may lack components of the Pol II holoenzyme (for example, Srb proteins, Gal11) that are functionally important *in vivo*^{21,22}.

A common view of the transcriptional activation process is that activator proteins stabilize the Pol II machinery at the promoter, thereby permitting increased transcriptional initiation²². In principle, activator proteins can interact with individual components of the Pol II machinery, and indeed, artificial connection of enhancer-bound proteins to TBP^{23,24}, TAFs (M. Keaveney and K.S., unpublished results) and components of the Pol II holoenzyme²⁵ can bypass the need for an activation domain. If natural activators interact with multiple components, individual components such as TAFs are likely to be non-essential for activation, even if they are potential targets. Thus, although it is possible to generate conditions in which TAFs are required for activation *in vitro*, they do not appear to be generally required *in vivo*. However,

at promoters lacking conventional TATA elements, which are inherently weak targets for TFIID, interactions of TAFs with basic transcription factors or with promoter DNA may be important for stabilizing the Pol II machinery. □

Methods

The parent strain ZMY60, containing copper-inducible alleles of *UBR1* and *ROX1*, was created as follows. A cassette containing the copper-inducible derivative of the *HIS3* promoter²⁶ and 2 kb of upstream flanking sequence was inserted at the initial ATG of a plasmid-borne genomic fragment of *ROX1* to create the *URA3* integrating plasmid ZM195. The same cassette was inserted at the initial ATG of *UBR1* to create ZM197. Both of these copper-driven alleles were introduced into yeast strain KY320¹¹ in successive two-step gene replacements. To create TAF disruption molecules, another cassette comprising an in-frame fusion of ubiquitin, arginine, LacI and the HA epitope driven by the *ANB1* promoter, was fused in-frame to a short 5' fragment of TAF coding sequence beginning at the initial TAF ATG. To create a given conditional knockout strain, the relevant TAF knockout molecule on a *URA3* integrating plasmid was linearized within the TAF coding sequence fragment and transformed into ZMY60. This integration results in homologous recombination at the TAF locus to yield a short, non-functional 5' TAF piece under its normal promoter, followed immediately downstream by a full-length copy of the tagged TAF under the *ANB1* promoter. Selection for uracil prototrophy was maintained in all experiments to avoid loss of the integrated plasmid. Details of the constructs and methods will be provided upon request.

RNA levels were determined by quantitative S1 analysis as described^{13,27}. All hybridization reactions contained at least two probes, such that the relative levels of all transcripts are internally controlled. The error for any particular RNA determination is $\pm 30\%$. TAF levels were determined by western blotting, with bands being detected by chemiluminescence (ECL). Relative levels of TAFs at various times were determined by comparing band intensities to serially diluted samples from wild-type cells.

Received 9 May; accepted 30 July 1996.

1. Tjian, R. & Maniatis, T. *Cell* **77**, 5–8 (1994).
2. Hernandez, N. *Genes Dev.* **7**, 1291–1308 (1993).
3. Struhl, K. *Annu. Rev. Genet.* **29**, 651–674 (1995).
4. Chen, J.-L., Attardi, L. D., Verrijzer, C. P., Yokomori, K. & Tjian, R. *Cell* **79**, 93–105 (1994).
5. Sauer, F., Hansen, S. K. & Tjian, R. *Science* **270**, 1783–1788 (1995).
6. Reese, J. C., Apone, L., Walker, S. S., Griffin, L. A. & Green, M. R. *Nature* **371**, 523–527 (1994).
7. Poon, D. et al. *Proc. Natl Acad. Sci. USA* **92**, 8224–8228 (1995).
8. Wang, E. H. & Tjian, R. *Science* **263**, 811–814 (1994).
9. Walker, S. S., Reese, J. C., Apone, L. M. & Green, M. R. *Nature* **385**, 185–188 (1996).
10. Chen, W. & Struhl, K. *EMBO J.* **4**, 3272–3280 (1985).
11. Iyer, V. & Struhl, K. *Mol. Cell. Biol.* **15**, 7059–7066 (1995).
12. Martens, J. A. & Brandl, C. J. *J. Biol. Chem.* **269**, 15661–15667 (1994).
13. Cormack, B. P., Strubin, M., Stargell, L. A. & Struhl, K. *Genes Dev.* **8**, 1335–1343 (1994).
14. Griggs, D. W. & Johnston, M. *Proc. Natl Acad. Sci. USA* **88**, 8597–8601 (1991).
15. Arndt, K. M., Ricupero-Hovasse, S. & Winston, F. *EMBO J.* **14**, 1490–1497 (1995).
16. Lee, M. & Struhl, K. *Mol. Cell. Biol.* **15**, 5461–5469 (1995).
17. Stargell, L. A. & Struhl, K. *Science* **269**, 75–78 (1995).
18. Koleske, A. J. & Young, R. A. *Nature* **368**, 466–469 (1994).

19. Kim, Y.-J., Bjorklund, S., Li, Y., Sayre, M. H. & Kornberg, R. D. *Cell* **77**, 599–608 (1994).
20. Workman, J. L., Taylor, I. C. A. & Kingston, R. E. *Cell* **64**, 533–544 (1991).
21. Koleske, A. J. & Young, R. A. *Trends Biochem. Sci.* **20**, 113–116 (1995).
22. Struhl, K. *Cell* **84**, 179–182 (1996).
23. Chatterjee, S. & Struhl, K. *Nature* **374**, 820–822 (1995).
24. Klages, N. & Strubin, M. *Nature* **374**, 822–823 (1995).
25. Barberis, A. et al. *Cell* **81**, 359–368 (1995).
26. Klein, C. & Struhl, K. *Science* **266**, 280–282 (1994).
27. Iyer, V. & Struhl, K. *Proc. Natl Acad. Sci. USA* **93**, 5208–5212 (1996).
28. Balasubramanian, B., Lowry, C. V. & Zitomer, R. S. *Mol. Cell. Biol.* **13**, 6071–6078 (1993).
29. Bartel, B., Wunning, I. & Varshavsky, A. *EMBO J.* **9**, 3179–3189 (1990).
30. Ray, B. L., White, C. I. & Haber, J. E. *Curr. Genet.* **20**, 25–31 (1991).

ACKNOWLEDGEMENTS. We thank J. Reese and M. Green for TAF antibodies, S. Buratowski for TBP antibodies, J. Haber for the isogenic tsm1-1 and TSM1 strains, V. Iyer and A. Campbell for technical advice, and L. Stargell, M. Keaveney, S. Buratowski, N. Kuldell, L. Apone, J. Reese and M. Green for discussion. Z.M. was supported by the Harvard-Markey Biomedical Scientists Training Program. This work was supported by grants to K.S. and P.A.W. from the NIH.

CORRESPONDENCE and requests for materials should be addressed to K.S. (e-mail: struhl@bcmp.med.harvard.edu).

YOURS TO HAVE AND TO HOLD BUT NOT TO COPY

The publication you are reading is protected by copyright law.

Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

ERRATUM

RNA-catalysed RNA polymerization using nucleoside triphosphates

Eric H. Eklund & David P. Bartel

Nature **382**, 373–376 (1996)

FIGURES 1*b–d*, 2*b–e* and 3*b* of this Letter were spoiled by a Moiré pattern, introduced as a result of an error during the production process. The figures are reproduced again here. □

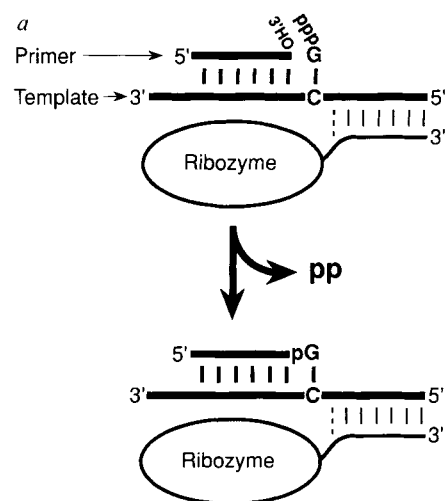


FIG. 1

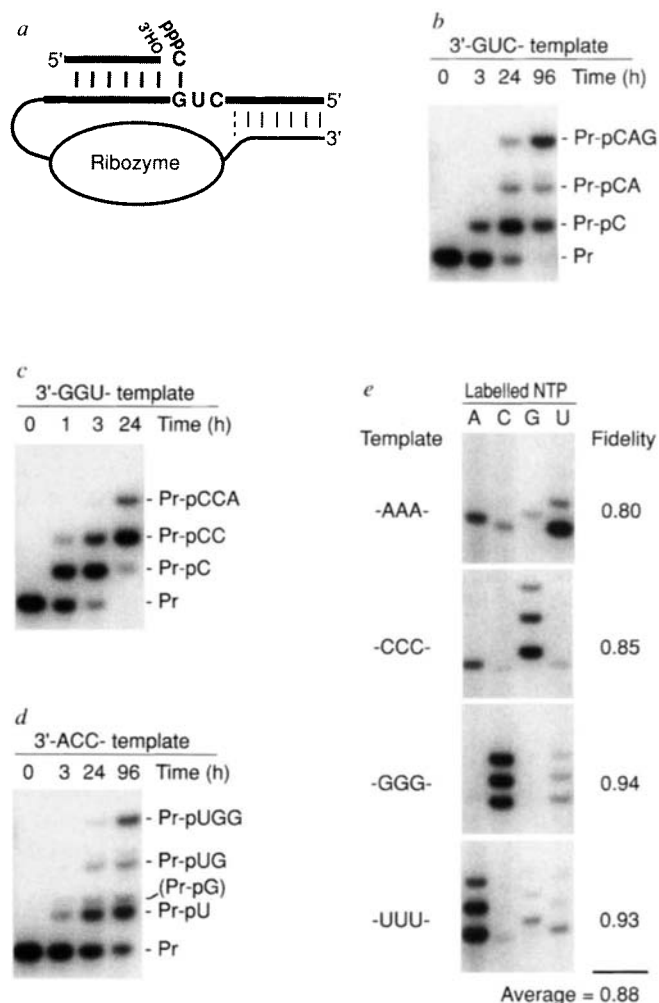
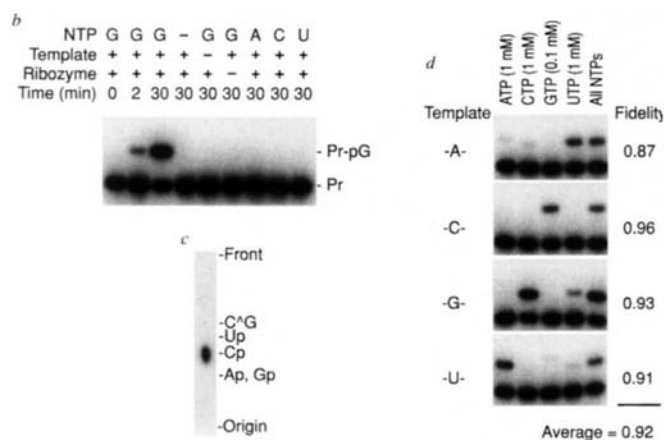


FIG. 2

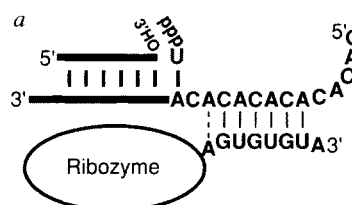
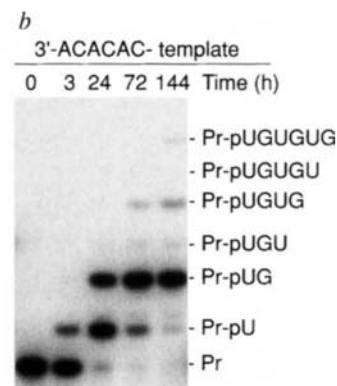


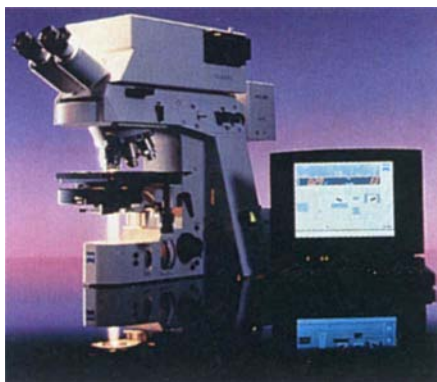
FIG. 3



Biomedical bulletin

Featured this week, a pot-pourri of useful equipment and reagents for biomedical research, including a mycoplasma removal agent, antibodies for apoptosis research, a biomedical imager and an oligonucleotide labelling kit.

A new **microscope series** from Carl Zeiss incorporates the latest 'CANbus' technology, allowing the microscope functions to be digitally controlled using a notebook or personal computer. The Axioplan 2 research microscope has digital stepper motors to drive the components and microprocessors built into the stand for control. Digital technology within the microscope stand allows settings to be 'memorized' for a specific technique and for a specific magnification. The correct adjustment of the microscope has been reduced to the activation of a single icon on the computer's graphical interface. The same principles apply to the new



CANbus technology from Carl Zeiss.

photomicroscope, Axiophot 2. Photomicrography is controlled within a software environment that can also store a wealth of distilled knowledge about photomicrographic techniques and interactively guide the user through the few steps needed to achieve the best result. The system is controlled by a personal computer, which with the software system, reduces photomicrography to a point-and-shoot process.

Reader Enquiry No. 100

Detect and destroy

Imperial Laboratories has announced the release of a **mycoplasma removal agent** called MC-210, which is designed to be effective in eliminating mycoplasma contaminants from cells in culture. The presence of mycoplasma in cultured cells is a serious problem, which may alter cell morphology, biochemistry and rate of growth. In some cases, the cells die because of this contamination, but often, the level of mycoplasma growth is slow, resulting in an ongoing low-level contam-

ination. This contamination may worsen when antibiotics such as penicillin, streptomycin and amphotericin B are used because they effectively control other microbial organisms but only inhibit mycoplasma without killing them outright. MC-210 has been shown to kill mycoplasma at a concentration of $0.5 \mu\text{g ml}^{-1}$ with an exposure time of only 7 days. It is stable at room temperature and is said to be easy to use: simply add it to the culture medium. According to the manufacturer, its effectiveness at low concentration and its short time of exposure to cells produce little or no cytotoxic effect on cultured cells. MC-210 is supplied in liquid form in a 5-ml pack, which contains 250 μg of product. This is said to be a suitable amount to treat 25 cultures in 25-cm^2 flasks on the basis of 10 ml of medium per flask, replaced once during the course of the 7-day treatment time.

Reader Enquiry No. 101

Serotec now offers a comprehensive range of **monoclonal and polyclonal antibodies for apoptosis research**. In addition to the common antibodies used by researchers, such as anti-Bcl-2, p53 and CD95(Fas), the company also supplies antibodies against emerging antigens of interest, including nuclear lamins, DNA-PKcs, CD40L, as well as other apoptosis-related antigens. These antibodies are suitable for use in a wide range of immunochemical techniques, including western blotting, immunoprecipitation and immunocytochemistry. They are supplied with comprehensive product information and can now be purchased preservative-free, which should prove useful for researchers using 'live' cell systems.

Reader Enquiry No. 102

A new ELISA for the quantitative **measurement of human apolipoprotein A (apoA)** in serum is now available from American Laboratory Products. ApoA is a glycoprotein linked to apolipoprotein B in the lipoprotein(a), Lp(a), particle. ApoA is made up of three different domains. The Kringle 4 domain is present in multiple copies, which vary by genetic code, resulting in multiple sizes of the Lp(a) particle. All isoforms of the Lp(a) particle can be detected with this new monoclonal antibody method. Lipemic, turbid

and even previously frozen samples can be measured with the apoA ELISA. According to the manufacturer, the method is precise and shows no cross-reactivity to either plasminogen or apoB. The kits are calibrated to a fully validated commercial Lp(a) preparation. One unit of apoA corresponds to approximately 0.7 mg of Lp(a) protein. The analytical detection limit of an undiluted sample is 0.05 U l^{-1} , with no observed 'hook effect' up to $9,600 \text{ U l}^{-1}$. This product is currently offered for research use only.

Reader Enquiry No. 103

Quantitative analysis

Spectronic Instruments has released the new Spectronic Genesys 2PC **spectrophotometer** with Windows95-based WinSpec software, which has been designed for faster collection, manipulation, and presentation of data and results. The new software offers the simplicity of a point-and-click environment and provides a full complement of applications: flexible data handling, a variety of file management options, multimedia and convenient output devices. The software incorporates advanced scanning, as well as calculations for standard curve, kinetics, absorbance ratio, absorbance difference, multiple wavelengths, and



Genesys 2PC spectrophotometer.

performance validation. The optical specifications of the Genesys 2PC are the same as the original spectrophotometer; however, this model was designed to be controlled by a PC through the WinSpec software or with end-user-compatible software. It is said to offer improved accuracy with 2-nm bandpass, increased sample throughput with an automatic eight-position cell holder and easy manipulation of scanning data with overlay, maths and derivative functions.

Reader Enquiry No. 104

MicroCal has developed a new generation of ultrasensitive differential scanning calorimeters for **biological calorimetry**. The VP-DSC micro-calorimeter uses recent developments from materials science, electronic and computers architecture to create a calorimeter that is easier to operate and has a significantly reduced footprint. It is stated to exhibit an increase in sensitivity over previous DSC models and requires one-tenth the amount of sample



MicroCal's VP-DSC micro-calorimeter.

material previously necessary. The calorimeter operates under VPViewer, designed for simple operation under a graphics-oriented, PC-driven user interface. Data analysis is performed using the MicroCal Origin software, which allows the user can customize the presentation of his or her findings within a publication-quality graphics environment.

Reader Enquiry No. 105

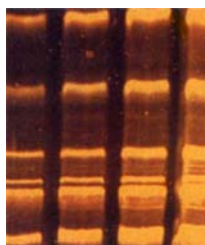
Miscellaneous

LaVision has introduced PicoStar, breaking new ground for the company with ultrafast gating **ICCD technology for biomedical imaging**. This product utilizes high-voltage pulsing technology to achieve gating times as low as 80 ps. The single-photon sensitivity, combined with high dynamic resolution, makes PicoStar suitable for time-resolved, two-dimensional imaging of laser-induced fluorescence, laser ablation, electron and energy transfer processes, photochemical reactions kinetics, biophotonics, neurobiology, biosensors, and signal propagation and transduction in tissues. The detector

head incorporates a photocathode, an image intensifier, a coupled fibre-optic slow-scan CCD and a TE cooler. The photocathode is sensitive over a wide spectral range from ultraviolet to near-infrared. The camera system comes with high-voltage pulsing electronics, allowing control over intensifier gain, gate width and delay. A comprehensive software package for image acquisition and processing makes the package complete for the spatial and temporal visualization, investigation and characterization of fast biophysical processes.

Reader Enquiry No. 106

Proteins can now be visualized after **gel electrophoresis** in one 30-min step using Sypro Orange from Bio-Rad. This fluorescent stain needs no destaining and is reusable, a significant cost-saving when compared with conventional staining methods. According to the manufacturer, the stain has a sensitivity equal to silver staining and does not give rise to overstaining, negative bands or variability problems. Stained proteins can be eluted or blotted and probed directly after visualization. The reagent does not stain DNA or RNA contaminants, but the compound stains proteins orange when viewed under ultraviolet light with absorbance peaks at 280 nm and 480 nm and emitting at 568 nm (orange). Excitation and detection are similar to gels stained with ethidium bromide, with excitation at the standard 302-nm ultraviolet wavelength. Gels can then be photographed or analysed using Bio-Rad's Gel Doc 1000 System.



Sypro Orange stain.

head incorporates a photocathode, an image intensifier, a coupled fibre-optic slow-scan CCD and a TE cooler. The photocathode is sensitive over a wide spectral range from ultraviolet to near-infrared. The camera system comes with high-voltage pulsing electronics, allowing control over intensifier gain, gate width and delay. A comprehensive software package for image acquisition and processing makes the package complete for the spatial and temporal visualization, investigation and characterization of fast biophysical processes.

Reader Enquiry No. 107

Cambio is now marketing a new kit from CPG called Label-It, which yields **labelled oligonucleotide probes** in 5 min

with high specific activity and an efficiency stated to be greater than 95 per cent. The reaction uses T4 polynucleotide kinase (PNK) to effect the transfer of the phosphate of ATP to the 5' terminus of DNA and RNA strands. The use of the PNK enzyme makes the system versatile and is said to be capable of labelling virtually any nucleic acid. There are two different reactions: the forward reaction labels nucleic acids containing a free 5'-hydroxyl, whereas the exchange reaction labels nucleic acids containing a 5'-phosphate with a relatively low specific activity. By using the forward reaction in association with dephosphorylation, it is possible to label nucleic acids with a 5'-phosphate of high specific activity. Single-stranded oligonucleotides can be labelled in just 5 min, with dsDNA labelling taking under an hour.

Reader Enquiry No. 108

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

PALL FILTRON

Two great names in Separations are now one!

For the full picture, visit <http://www.pall.com>
 For a catalogue fax +44 1705 302510
 Address specific enquiries to
Ian Sellick@pall.com, or write to
Pall Filtron BioSeparations Division,
 Pall Europe Ltd. Europa House Havant Street,
 Portsmouth PO1 3PD, UK. Visit Pall on
Stand 215 at BioInnovations,
 Wembley, London September 24th-26th 1996.

Filtration • Separation • Purification

READER ENQUIRY NO. 75

LIPOSOMES

REPRODUCIBLE - HOMOGENEOUS -
EXTREMELY RAPID

NEW

Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

P.O. Box 650126, D-81215 Munich
 For USA call 1-800 Liposome

READER ENQUIRY NO. 62

NEWCASTLE UNIVERSITY
 MOLECULAR BIOLOGY UNIT

PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For further information contact Dr. Joe Gray
 Phone: +44 191 222 8612
 Fax: +44 191 222 8129
 Email: JOE.GRAY@newcastle.ac.uk

READER ENQUIRY NO. 69

DNA SEQUENCING

£45 per reaction

- PCR products
- M13 inserts
- plasmid inserts
- single strand DNA
- double strand DNA
- no result? no charge!

By research labs for research labs

Molecular Biology Services
 King's College School of Medicine and Dentistry
 Rayne Institute, 123 Coldharbour Lane, London SE5 9NU
 Tel: 0171-346 3126 Fax: 0171-733 3877

READER ENQUIRY NO. 68

The world of the postdoc

Opportunities abound for those who wish to make scientific research their career, if a nomadic lifestyle and short-term positions are acceptable. Advice abounds, particularly on the Internet. *Nature's* global review begins with the biggest postdoc jungle of them all.

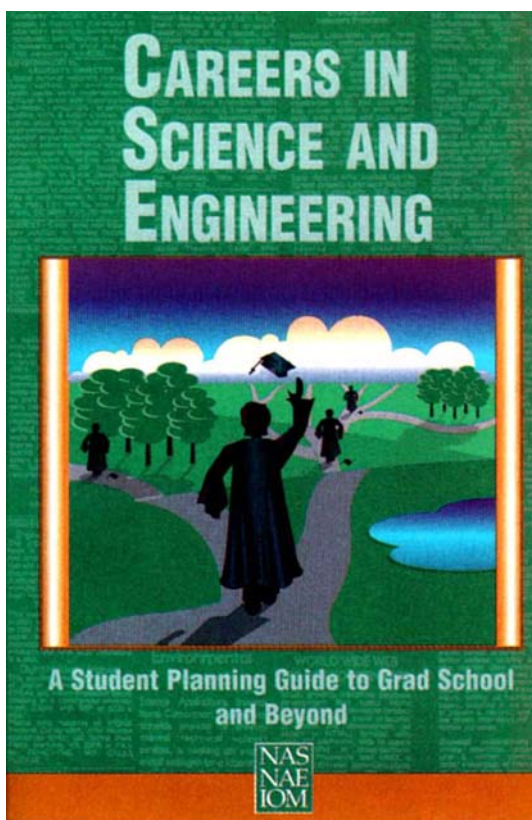
Washington. A report issued earlier this year by a committee of the National Academy of Sciences (NAS), *Careers in Science and Engineering, A Student Guide to Grad School and Beyond**, found that "more than half the students who have received PhDs in science and engineering obtain work outside academic research — a proportion that has increased steadily for two decades". The Academy's Committee on Science, Engineering and Public Policy (COSEPUP) found that fewer than half of PhDs in academic institutions hold tenure-track positions and almost half of these are non-research positions (teaching, for example). Given that fewer than a third of science and engineering PhDs obtain a tenure-track position within 5–8 years of receiving their degree, many young scientists, willingly or not, are having to explore alternative career paths.

Many feel that postdoctoral fellows are the worst off. With fewer postdocs finding permanent employment, and a steady stream of graduates entering the job market, there is a log-jam at the postdoctoral level. According to a survey carried out by the American Chemical Society in the summer of 1992, the percentage of applicants accepting postdoctoral positions in 1992 was 43 per cent, up from 34 per cent and 37 per cent in 1990 and 1991, respectively. In 1993, *Chemical Engineering News* described the situation as a "clogged pipeline", where new (chemistry) PhDs were accepting postdoc positions they would not ordinarily have considered. In the book *Rethinking Science as a Career**, Sheila Tobias, Daryl Chubin and Kevin Aylesworth reported to Research Corporation, a foundation for the advancement of science based in Tucson, Arizona, that the postdoctoral fellowship can become a "temporary holding pattern" when permanent employment is not available or when postdoctoral terms are extended to keep research costs down. The result is that PhDs end up doing two or more successive fellowships and their salaries fall further behind because of this. In the book the authors give their explanation as to how the current glut of young scientists has developed and provide information

on alternative scientific degrees, flexible career paths and how the demand for scientists could be restructured. The authors go on to suggest that the sense of betrayal now being felt by many young scientists stems (at least in part) from a belief that

these career investments, Kasanicki was having second thoughts about research but wanted to give it another chance by accepting a postdoctoral position at Boston University. Realizing that "to succeed in research you have to love it", she attended a symposium on alternative careers in science, which brought her to a crossroads in her own career. A summer course gave her insight into some of the legal issues affecting high-technology industries, and confirmed that her interest lay more in the commercial development of research. Upon returning to the United Kingdom, Kasanicki joined the patent department of British Biotech. This was followed by legal training at the law firm of Simmons and Simmons in London, where she worked in areas of intellectual property and corporate finance. She now finds that her knowledge and scientific skills, such as analytical thinking, technical writing and understanding and evaluating scientific information, are important in her current employment. Kasanicki recommends that the training of young scientists should place greater emphasis on fostering skills such as technical writing, giving scientific presentations, statistical analysis or other areas where an individual may need greater proficiency.

According to COSEPUP, more than 14 per cent of the companies recruiting at the Massachusetts Institute of Technology in 1995 were financial companies, almost three times as many as in 1983. Chris Finger received his PhD in June 1995 from the Program in Applied and Computational Mathematics at Princeton University and started working in January this year in the risk-management group of a large bank in New York. His dissertation was on stochastic processes, which is somewhat applicable to his current work, although all his studies did not focus on finance. Finger says that he chose finance as a career alternative to academic research for a variety of reasons:



Career paths — advice for young scientists.

they have not been prepared practically or psychologically for work on the outside.

A new outlook may be necessary. COSEPUP recommends that young scientists approach their careers with a broad view, with the well-developed set of professional survival skills that today's job market requires. As an example, it states that multidisciplinary research experience or dual masters' degrees are often more attractive to potential employers than a narrowly focused PhD.

Young scientists are also now discovering that advanced degrees in science can make important contributions in many more areas than just research. For example, Mary Kasanicki began her scientific study with a first degree in biochemistry and followed that up with a PhD in biochemistry from University College London. Even after

Cover of *Careers in Science and Engineering*

On other pages

- NIH helps with loan repayments 197
- Situation in Europe 199
- Positions galore in Japan 200

* *Careers in Science and Engineering, A Student Guide to Grad School and Beyond*, National Academy Press, 2101 Constitution Avenue, NW, Lockbox 285, Washington DC 20055, USA. *Rethinking Science as a Career*, Research Corporation, 101 North Wilmot Road, Suite 250, Tucson, Arizona 85711-3332 USA.

the lack of a real-world connection to his academic work, the lack of a collaborative work atmosphere, the timescale of projects in academic research (he wanted a project that ended at a definable point) and the mediocre research job market.

His present research group consists of five or so other PhDs of varying kinds (in mathematics, physics, economics and engineering). "The group's work has allowed me to be exposed to a variety of financial products, and all of us bring expertise in different areas to the table", Finger says. "One of the biggest challenges that did not exist in an academic research environment is the need for me to communicate ideas and issues to people with vastly different backgrounds from my own." Getting over this 'cultural barrier' often takes quite a bit of creativity, he adds.

Institutional adaptation

Has the training of PhDs adjusted to the demands of the job market? In recent years, institutions have begun to change, especially with regard to the postdoc. Fellowship networks, such as those at the National Institutes of Health (NIH), Johns Hopkins, the Mayo Clinic Graduate School and the University of California at San Francisco, are helping postdocs consolidate their interests on-campus. According to Michael Fordis, director of NIH's Office of Education, the NIH Postdoctoral Fellows Committee had its beginnings about four years ago and has been "a real asset to NIH". The committee was structured to be a grassroots organization, allowing fellows to have input into their own training. It sponsors workshops on career development, organizes programmes to help fellows to establish pathways to tenure and hosts job fairs to bring fellows into contact with industry representatives. The following topics, for example, will be addressed in a forthcoming seminar series jointly sponsored by the fellows' committee, the NIH Office of Research on Women's Health and the NIH Office of Education: finding employment, negotiating a job offer, life as a professional, making oral presentations, obtaining research funding, writing and publishing, and teaching. The fellows' committee has also developed an awards programme called the Fellowship Award for Research Excellence (FARE), which provides money to fellows for travel to scientific meetings to present research. In March this year, a joint committee of fellows and faculty members granted 90 FARE awards.

The Office of Education also supports an Internet site for the fellows' committee, which includes the fellows' handbook, a discussion network where questions can be posted and where fellows and faculty can exchange ideas. The site includes details of the fellows' activities, electronic applications for FARE awards, industry job listings and links to journals, such as *Nature*, *Science* and the *New England Journal of Medicine*, as well as links to federal employment databases.

TABLE 1 INTERNET RESOURCES FOR MATERIALS USED IN WRITING THIS REVIEW

Online job listings and career information	http://
American Chemical Society	www.chemcenter.org/profservices.html
American Institute of Physics	www.aip.org/careersvc/
Bio Online	cns.bio.com/bio.html
The Chronicle of Higher Education	chronicle.merit.edu/ads/links.html
FASEB	www.faseb.org/default.html
National Assoc. of Colleges and Employers	www.jobweb.org/
NAS careers and research opportunities	www2.nas.edu/wwwcat/Careers.html
National Institutes of Health	www.nih.gov/
Funding	
National Academy of Sciences	www2.nas.edu/wwwcat/Fellowship.html
Research Corporation	email: awards@rescorp.org
NSF grants/funding/programme areas	www.nsf.gov/nsf/homepage/grants.htm
NSF Publications and Data	www.nsf.gov/sbe/srs/pubdata2.htm#Topic 3
Postdoctoral networking	
Young Scientist Network	www.physics.uiuc.edu/ysn/
Industry listings	
Bristol-Myers Squibb	www.bms.com
Eli Lilly (postdoc programme information)	www.lilly.com
Genentech	www.gene.com
On-line supervision	
Networking the Network (P. Agre)	weber.ucsd.edu/~pagre/network.html
Search Masters International (D. Jensen)	smi.bio.com/
NAS Career Planning Center	www2.nas.edu/cpc/index.html
Sites through Nature	
Wyeth Ayrest	www.nature.com or
National Institutes of Health	www.america.nature.com

NIH fellows can also post their *curricula vitae* (CVs) electronically on this site.

Importance of networking

In an interview with Richard Bolles, author of the book *What Color is Your Parachute?* (Ten Speed Press, PO Box 7123, Berkeley, California 94707) David Jensen, a recruiter from Search Masters International in Sedona, Arizona, reports one case where a job advertisement received a total of 1,470 applicants, with only one candidate getting a job. With odds like these, Jensen says "the nature of the job market is — and always has been — the business of making quality contacts". In his extensive on-line career centre (see Table 1), Jensen reports the findings of the outplacement firm of Drake Beam Morin, which found that only 8 per cent of its clients obtained jobs by mass mailing a cover letter and resumé/CV, 9 per cent in response to job advertisements, 15 per cent with the firm's help and 68 per cent through a network of professional contacts. "Don't believe that CVs alone will do the job", he says. Jensen also recommends that people respond to job advertisements, but only to those that "fit like a glove".

Philip Agre (assistant professor in the Department of Communication, University of California, San Diego) author of the article *Networking on the Network* (see Table 1), feels that young scientists need to understand the professional persona of the people in their network, an important step in developing a strong professional presence of their own. One exercise in self-promotion, which Agre recommends, is for young scientists to develop their professional persona by creating a home page. He says, however, that "the Internet is only part of a larger picture: add it to the list". There are a host of elements that go into a professional persona besides the merits of one's scientific work.

These can include letter writing, scientific publication and presentation, collaboration, leadership and teamwork skills, and consultation and supervising.

Another area that can be useful for professional advancement and network building is developing organizational skills. Agre recommends organizing a journal club or conference session, urging young scientists to understand that these exercises are not self-sacrifice without reward. Networking can also take the form of reciprocal professional courtesies, such as the exchange of pre-publication drafts. The process improves the work of the participant through *quid pro quo* commentary. Agre says that "this draft-exchange ritual is tremendously important, but nobody ever seems to teach you how it's done. When in doubt, ask for help." On Agre's Web site, he discusses the etiquette of professional reciprocity and how to approach people who might become part of a network, especially peers who are at a similar stage of career development. The site also has a reference list, which should serve as a general guide to literature on professional networking.

Another good source for material in the area of professional development is the reference list of the COSEPUP report (see Table 1). The information provided in this book pertains to survival skills that are useful in all areas of science and encourages the development of skills such as large-scale management, executive decision-making, technical writing, public speaking, teaching, supervising students, and teamwork and leadership skills. There is an ever-increasing number of sources for this kind of professional resource. With regard to supervision, for example, the American Institute for Physics is looking for volunteers to participate in its on-line supervision programme (see Table 1 for AIP Web site).

Faculty, department, institute and corporate home pages can provide a useful way of obtaining background information on the people that should be included as professional 'centres of influence', a term used by Jensen. The centres of influence are the people who can tip the balance in the search for a position by making a recommendation, an introduction, suggesting a contact or alerting young scientists to the most appropriate time and place to publish their work.

Someone who is considering a career in industry, for instance, might do well to start with corporate Web sites (see Table 2). Eli Lilly's site, for example, includes information about the company's postdoctoral training programme. According to Joreg Pfeifer, Lilly's director of research personnel, the company plans almost to double its number of postdoctoral research positions (from 40 to 70) in the next year. The company will be hiring across-the-board in up to 12 disciplines. The appointments will be for one year, renewable for up to a maximum of three years. Jensen likens networking to "building a pyramid from the top down". The top stone in the pyramid is often the undergraduate or graduate adviser.

Scientific meetings remain the best arena for developing contacts. Preparation is the key to successful network development at meetings. Compiling a list of potential contacts beforehand, with background information about each, is recommended. At meetings, Agre stresses a low-key approach such as asking a potential contact for advice, an opinion, additional contacts in the field or simply using the opportunity to discuss a mutual area of research interest. A direct question about a possible job does not gen-

erally meet with success; however, asking a well-connected person for advice about job hunting is both good and normal, he says. The best place to prepare for meetings with potential contacts, according to Agre, is the library. Begin by compiling a list of contacts and then learn more about them.

Jensen also advises that the best networking is rooted in the information-gath-

ering process. Unfortunately, scientists often rely solely their CV, thinking their science will sell itself, he says. According to Jensen, only about 15–20 per cent of all open positions are advertised. "When you limit your job search to 20 per cent of the market you place unnecessary constraints on the already short supply of potential opportunities." **Brendan Horton**

NIH lends a hand with loan repayment

Houston. For Claire Esposito, a critical-care fellow in the Clinical Center at the US National Institutes of Health (NIH), the agency's student loan repayment policy seemed too good to be true. As it turned out there was no catch. Esposito, now in her second year of a research fellowship and with student debts that would cause many to lose sleep at night, says participation in NIH's loan repayment program (LRP)* has "relieved a significant financial stress", the monkey on one's back that follows one through medical school, residency training and afterwards. Esposito says she wanted to "learn the business of research" but was originally reluctant to embark on a two-year fellowship because it would mean two more years of struggling to get by on a training salary. She admits that the LRP changed her outlook and she now intends to stay at NIH beyond her original two-year commitment.

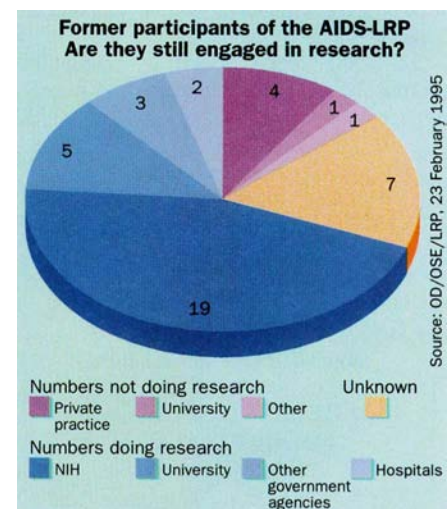
So what level of student debt is commonplace? According to the Association of American Medical Colleges (AAMC), the average student debt on graduating from medical school, for example, is about \$70,000. But Marc S. Horowitz, director of NIH's Office of Loan Repayment, believes this figure is "deceptively low" because it fails to take into account the additional interest accrued on those student loans during residencies and internships. At NIH, "we see candidates with \$120,000 and \$130,000 in debts".

With debts like these it is not hard to see why so many individuals (particularly MDs) find they have to moonlight to make ends meet early on in their careers. Many take the path of least resistance and are lured into higher paying jobs in the private sector. A neurologist employed in a basic science laboratory at NIH can expect to earn between \$45,000 and \$55,000, says Eugene O. Major, chief of the laboratory of molecular medicine and neuroscience in the National Institute of Neurological Disorders and Stroke. On the outside, if they were fully board-certified, "their income would be doubled".

Faced with salaries out of step with the private sector but a need to intensify its

research efforts in the AIDS/HIV area, NIH implemented its first LRP in 1989 in the hope that the programme would provide the incentive necessary to attract bright young researchers into its intramural AIDS research programme. So far, more than 90 researchers (primarily MD and PhDs) have benefited, with almost \$5 million repaid in student debts.

Robert Yarchoan, a section head and senior investigator at the National Cancer Institute, has shepherded at least four people through the AIDS LRP over the years (all of whom are still involved in some



aspect of AIDS research, though not necessarily at NIH). He says it is difficult to say with certainty that the programme has increased the cadre of investigators trained in AIDS research, although he suspects it has. Where he has found it particularly useful is in recruiting and retaining MDs, as it helps reduce their financial anxiety from large, outstanding medical school loans during the first couple of years of research, when many of them may otherwise get discouraged and go into private practice. Major's experience with the programme has been much the same. Having had three people participate in the LRP, and one about to apply, he believes it has enabled him to recruit good people to his laboratory and to keep people who are good for longer.

One measure of the success of the initiative is the extent to which it encourages people to stay in research — whether at NIH or elsewhere. A follow-up survey of 42 former AIDS research LRP beneficiaries (see pie

TABLE 2 CORPORATE WEB SITES

ALSAC-St Jude's Children's Research Hospital http://www.stjude.org
Bayer Corporation http://www.bayer.com/english/0000home.htm
Biotechnology Industry Organization http://www.bio.org
Bristol-Myers Squibb Company http://www.bms.com
Ciba Pharmaceuticals http://www.ciba.com
Eli Lilly and Company http://www.lilly.com
Glaxo Wellcome Inc. http://www.glaxowellcome.co.uk
Hoffmann-La Roche Inc. http://www.roche.com
Johnson & Johnson http://www.jnj.com
Pharmaceutical Research & Manufacturers of America http://www.phrma.org
Pharmacia and Upjohn Inc. http://www.pharmacia.se
Procter and Gamble Health Care http://www.pg.com
Sandoz Pharmaceuticals http://www.sandoz.com/
Warner-Lambert Company http://www.warner-lambert.com

* For more information on NIH's LRPs, contact the Office of Loan Repayment, tel. 1-800-528-7689, fax 301-480-5481, e-mail: mh18k@nih.gov. (A list of state and other loan repayment initiatives can be obtained from AAMC, Washington, DC, tel. 202-828-0681, fax 202-828-1125.)

chart) showed that of the 35 individuals who could be contacted, 29 are still in research, and of those, 25 are still involved in AIDS-related research.

The loan repayment initiative was expanded in 1994 to include an LRP for clinical researchers from 'disadvantaged' backgrounds — a move made (in part) to help redress the balance between the number of basic and clinical researchers at NIH. Horowitz is keen to emphasize, however, that the programme is not solely targeting under-represented minorities: individuals from low-income families, as well as those with 'special' personal circumstances, are eligible for consideration. Twenty-two individuals have benefited so far.

Just because a position qualifies the holder for participation in an LRP does not mean the cheque is in the post. The awards are 'competitive', as are renewals, although about 70 per cent of those applying to the AIDS research LRP, and so far all of those applying to the clinical research programme, are successful, says Horowitz. In making its decision, the review committee pays close attention to the kind of supervision and training an individual will receive while at NIH.

To be eligible for consideration, individuals must be US citizens, nationals or permanent residents and must owe more than 20

TABLE 3 NIH LOAN REPAYMENT PROGRAMS (LRPs)

Programme	Targeted at	Qualifying areas of research	Employment requirement	Amount of repayment	Tax liability reimbursement benefits	Renewal option
NIH AIDS research LRP	Physicians, scientists and nurses or equivalent degrees engaged in qualified HIV/AIDS research	Research on the aetiological agent, pathogenesis, therapeutics, vaccine development, behavioural intervention, epidemiology and natural history of HIV infection	2-year minimum	Up to \$20,000 per year for 2 years	Yes	Yes
NIH clinical research LRP for individuals from 'disadvantaged' backgrounds	Clinical researchers with MD, DO, DDS, DMD, ADN/BSN or equivalent degrees from 'disadvantaged' backgrounds	Biomedical and behavioural studies of the aetiology, epidemiology, prevention (and prevention strategies), diagnosis or treatment of diseases, disorders or conditions, including clinical trials	2-year minimum	Up to \$20,000 per year for 2 years	No	Yes

per cent of their annual NIH salary (see Table 3). Both programmes offer a maximum annual loan repayment of \$20,000, with the amount based in part on the proportion of an individual's qualifying debt relative to their NIH salary. In return, beneficiaries are required to commit themselves to at least a two-year stint at NIH. With the AIDS research LRP, NIH will also meet any increased tax liability that results, as payments made to a beneficiary are considered as 'income'. This is not currently offered with the clinical research LRP, although there is a move afoot to fix this inequity.

Whereas the existing programmes were implemented to respond to a specific identi-

fied need, Horowitz says that what NIH really needs now "is not a proliferation of disease-specific LRPs but a 'general' loan repayment authority" for tenure-track researchers, senior staff fellows and others who are not eligible for the existing programmes and who are prepared to spend at least three years at NIH. A general LRP was authorized by Congress in 1993, but a lack of available funds has so far prevented any progress on this front. If the appropriation is forthcoming for fiscal year 1997 (which begins on 1 October), and Horowitz is quietly optimistic, then NIH hopes to be able to fund between 15 and 20 individuals in the first year.

Diane Gershon

ADVERTISEMENT

SYNTHETIC VIRAL PROTEINS

HIV-1 TAT HB10 (1000 USD/100 µg)

HBV pre-S1 adw (1000 USD/100 µg)

myristil-HBV pre-S1 adw
(800 USD/50 µg)

AND RELATED ANTIBODIES

Fax +39 823 612222

Email info@tecnogen.it

TECNOGEN SCpA - Science Park
81015 Piana di M.Verna - CE - Italy

WEB Site: <http://www.tecnogen.it>

READER ENQUIRY NO. 80

The Danish Centre for Human Genome Research

is pleased to announce the release of the
Human 2-D PAGE Databases
for Proteome Analysis



<http://biobase.dk/cgi-bin/celis>

READER ENQUIRY NO. 78

NIH biomedical foundation finds its feet

WORTH keeping an eye on in the coming months is the new National Foundation for Biomedical Research, now up and running having recently been incorporated in the State of Maryland as a tax-exempt (not-for-profit) foundation. The foundation hopes to raise money through donations from private citizens, corporations and philanthropic foundations to support special projects that fall within the mission of the National Institutes of Health (NIH) but which NIH is unable to fund adequately — or at all. Although the foundation was authorized by Congress in 1990 (and re-authorized in 1993) following a recommendation by the Institute of Medicine, Congress has yet to appropriate funds to cover its administrative costs.

It may be early days, but one of the areas where the foundation will probably focus its efforts and resources in the near term is in supporting research training, such as fellowships at the graduate and senior level — both at NIH and within the extramural community. Providing fund-

ing for researchers to take sabbaticals in other laboratories, is another. The foundation also plans to support public education programmes, as well as to organize regional seminars and meetings that will bring primary health-care providers up to date with the latest research findings.

George J. Galasso, former associate director for extramural affairs at NIH and now the foundation's executive director, is busy putting flesh on the bones of the organization. Galasso, who retired from NIH last January after 28 years, says that although the foundation will seek input from NIH, it will operate independently, under the guidance of a board of directors made up of prominent academics and corporate leaders. Nobel laureate Paul Berg, director of the Beckman Center at Stanford University School of Medicine, will serve as acting chairman until a permanent chairman is appointed. The nine-member board (there are two vacancies) will meet for the first time this month.

D. G.

The view from Europe

Although the prospects for European postdocs are as uncertain as anywhere, young people are flocking to science as a career.

Munich. It is not difficult to find a postdoctoral position in Europe at the moment. The question is: is taking a postdoc a sensible career move? The unpleasant truth is that the time when completion of one or two postdoctoral fellowships meant a strong likelihood of a safe academic career, with the tantalizing prospect of a full professorship for the most successful, has long since passed. The recession has seen to that.

Talk to any European postdoc, and you will come away with two impressions: enthusiasm verging on passion for the research

itself; and dissatisfaction verging on bitterness for the lack of certainty about the future. On top of this there is weariness with the continual need to seek the next short-term contract, and a lingering suspicion that postdocs are being exploited.

There is wide concern about the lack of career advice, and the poor supply of permanent jobs, on offer either in academic institutions or in industry. And everyone is aware that the issue is time-sensitive: having a postdoctoral position beyond one's 30s reduces the chances of finding a suitable permanent job.

Nonetheless, young European scientists choose the nomadic way of postdoc life in droves. Most enter idealistically, with high hopes for the future. Some, it must be said, see the life as an alternative to unemployment, albeit a positive one. In Germany, for example, the number of applications for fellowships is rising. The chances of success with the largest grant giver, the Deutsche Forschungsgemeinschaft (DFG), remain, at 70 per cent, very high.

The total number of postdoctoral fellows working in Germany is difficult to assess. As in the United Kingdom and France, there is no central body with responsibility for keeping a check, or for following up on what happens to postdocs as they leave the system.

The DFG funds two types of fellowship. It offers around 1,500 personal fellowships each year, mostly for a two-year period. Young scientists themselves may apply for these, but the fellowships come with no social benefits such as pension or health insurance. More coveted are the 7,000 or so 'on-staff' university positions, and a further 5,000 or so postdoc positions at non-university research institutes such as those belonging to the Max Planck Society, which are paid on the standard academic scale and include full social benefits.

But these positions carry an obligation to teach. And the candidate may not apply for a position directly but must depend on the patronage of the scientist in charge of the relevant research project. About 2,300 additional postdoc positions are funded indirectly through DFG grants to senior scientists who may then recruit young scientists themselves. Smaller grant-giving bodies such as the Volkswagen Foundation also operate in this way. The DFG is planning a small-scale new scheme to open up at least some of these positions to direct applications.

The DFG has been accused by the government, which holds its purse-strings, of misplanning its allocation of fellowships by not taking into account the economic

demand for particular subjects. Bruno Zimmermann, a spokesman for the DFG, defends the decision-making process which takes only scientific quality into account. "We are not a tool of educational politics", he says.

The exact numbers of postdocs in the United Kingdom are not officially collected and published, but a study conducted by the House of Lords Select Committee on Science and Technology last year (see *Nature* 376, 715; 1996) put the figure at around 20,000. The report strongly criticizes the poor employment conditions that these scientists are forced to accept by their employers: low pay, no pension provision or other social benefits such as above-statutory maternity pay, and no career structure. According to the report, many of the best scientists are thus leaving the UK academic research system and seeking better condi-

INTERVIEW

MARIEL Donzeau obtained her doctorate degree at the University of Bordeaux in France at the end of 1993. She hopes eventually to get a *maitre de conference* position (corresponding to an associate professor) in France and therefore wanted some years' research experience as a postdoc.



Mariel Donzeau: Retaining scientific contacts is very important.

Lack of adequate national support forced her to look for fellowships abroad, something that turned out to be not too difficult. She was able to pick and choose from around ten suitable vacancies.

She chose eventually to join Ernst-Ludwig Winnaker in Munich, primarily because the project on offer — to establish a yeast-based cloning technique to study the interactions between human viruses and their target genes — matched exactly her experience and interests.

The grant was provided through the European Commission's Human Capital and Mobility (HCM) Programme. It was also astonishingly generous, offering more than DM5,800 per month tax-free. But future applicants will not be so lucky. The HCM programme expired last year and its successor, the Training and Mobility of Researchers (TMR) programme, now offers so-called Marie Curie Fellowships that are more in line with local rates for postdoctoral fellowships.

Donzeau preferred to stay in Europe rather than going to the United States because she recognized the need to stay in personal contact with French researchers. She learnt this lesson from colleagues who, after some years in the United States, experienced great difficulty in finding a permanent position in their home country.

G. S.

INTERVIEW

ANDREAS Gellrich is working on his second postdoc at DESY, the large synchrotron facility in Hamburg. He is part of the HERA B team, preparing an experiment with DESY's hadron-electron ring accelerator to study fundamental questions in particle physics concerning subatomic symmetry (CP-violation). His job is to establish a 'computer farm', by synchronizing around 200 personal computers to evaluate real-time data from 10 different subdetectors, in experiment. He hopes that this experience will eventually help him to find a job in industry.

Gellrich completed his doctoral thesis for the University of Hamburg in May 1994 at DESY, and was then given a further six-month postdoctoral contract to continue his research and decide what to do next. During this time he applied without success for a wide range of industrial positions. He was disappointed to find that few companies were interested in applicants without direct experience in the exact, highly specialized technical areas specified in the job description. He was amazed, he says, to find that companies were not receptive to his view that a general education in physics gives one the flexibility to tackle any such problem.

Aware that his career-clock was starting to tick loudly — his colleagues had warned him that it is hard to find a permanent job after one has turned 40 — he decided that he would fill in a couple more years with a further postdoc. There were plenty to choose from, but he decided to stay at DESY and continue his current work.

But he is dissatisfied, and complains about the total lack of formal career advice for postdocs in Germany. At 33 he has no idea what his future will bring. The human cost of this uncertainty means that he has delayed plans to start a family. The system discourages the best young scientists from staying in the university system, he complains. If nothing is done to protect young researchers, "the politicians will one day wake up and find that Germany has lost a generation of scientists".

G. S.

tions and higher status abroad or in industry.

The situation is so acute that the concept of a postdoctoral fellowship — a short training period for a scientist with a PhD — has to a large extent fallen by the wayside. The terms 'short-term contract staff' or 'contract researchers' have become common currency, reflecting the fact that being a postdoc is now seen as a way of life. So it is no surprise to find that two-thirds of research staff in biochemistry and clinical sciences in the United Kingdom, and half of all researchers in physics, are on short-term contracts.

Nevertheless, according to Philip Diamond of the Institute of Physics (IOP), the number of postdocs in physics has grown at least fourfold in the past few years. A recent informal study conducted by the IOP revealed a vociferous crowd of contract researchers complaining of exploitation and expressing the fear that they will be 'dumped' when they become older and more expensive to keep on. Postgraduate students, however, expressed determination to continue in their own research area 'come hell or high water'. What is missing from the system, says Diamond, is career advice, urgently needed before scientists reach the watershed of their 30s.

Contract researchers in the United Kingdom have one piece of good news. A new concordat agreed between universities and grant agencies gives better employment conditions, such as recognition of previous experience when setting salary levels. The concordat will be fully launched later this month.

The United Kingdom differs from other European countries in that there are many more different sources of fellowship funding, including many charities (foremost among them the Wellcome Trust), as well as the national research councils.

The situation in France is particularly tough. In 1995, 10,000 young scientists completed their PhD theses, and this year the figure is expected to be as high as 13,000. Those wishing to stay in academic research will have to fight for a total of 2,500 permanent research assistant positions a year at the CNRS (the chief French research organization) or a post as associate professor at a university.

Those wishing to go into industry will find help from the non-profit-making Bernard Gregory Association. Its director, Marc Joucla, says that the association arranges around 400 industrial positions each year to scientists with PhDs. But he warns that too many years of postdoc experience have a very negative effect on

chances of a job with industry. It is important not to let too much time elapse after getting one's PhD, he says.

There are virtually no postdoctoral funding programmes in France — fewer than 400 French postdocs work in their home country at any time — so scientists have to look abroad for opportunities. An estimated 10,000 young French scientists are in postdoctoral positions abroad. Most report difficulties in returning home later, as they say that personal contacts are much more important than they should be for senior appointments, so people who have worked abroad are at a disadvantage.

It is interesting that the United States is no longer the favourite destination for European postdocs. Although 55 per cent of travelling postdocs went to the United States in 1992, by 1994 this number had fallen to 43 per cent. Over the same time period, the European Union emerged as the favourite destination, now hosting 45 per cent of postdocs compared to 30 per cent in 1992. The European Commission's Human Capital and Mobility (THR) Programme, which has now been succeeded by

the Training and Mobility of Researchers Programme, has been immensely popular as a means of getting postdoc experience in another European country. The commission has recently relaunched its postdoc programme under the name Marie Curie Fellowships. Last year it received nearly 25,000 applications; 850 fellowships were granted. Similarly the European Molecular Biology Organisation in Heidelberg, which offers around 150 fellowships a year, attracts nearly twice as many applicants now (around 900) as it did 5 years ago. The number of awards unfortunately has not significantly changed.

The problems, and the accompanying lack of information about employment prospects in France, have spawned something unique in Europe — self-help groups. A group of PhD students called 'hotdocs' communicate news and information over the Internet. Postdocs can be helped by the group called ANDES (Association Nationale des Docteurs Ès Sciences) which publishes an annual guide to finding fellowships and postdoctoral positions abroad.

Alison Abbott & Gabor Stiegler

Postdoctoral positions galore in Japan

Tokyo. Opportunities for postdoctoral research in Japan for both Japanese and non-Japanese scientists will increase dramatically over the next five years under a 5-year government plan to promote science and technology. The plan, approved by the Japanese cabinet in June under a new law passed late last year (*Nature* 378, 227; 1995), requires the government to double the number of domestic postdoctoral fellowships for young Japanese scientists to about 10,000 by the year 2000. Similarly, the number of fellowships for non-Japanese scientists offered under two government schemes will increase dramatically from the present level of 760 a year to 2,050.

Both the Japan Society for Promotion of Science (JSPS) and the Science and Technology Agency (STA) offer postdoctoral fellowships for non-Japanese. Introduced in the 1980s, the schemes are open to scientists throughout the world and offer generous tax-free salaries and accommodation allowances (see *Nature* 362, 867; 1993). At present, JSPS offers 420 fellowships a year for research in universities and university-affiliated institutes, while the STA provides 340 for research in national research institutes and other government-related research laboratories. The JSPS scheme will be expanded to 1,050 people and that of the STA to about 1,000 over the next 5 years.

The Japanese government has in the past experienced difficulty in filling all the positions, partly because of poor promo-

tion of and publicity for the schemes and partly because Japan does not offer a particularly appealing research environment, especially for researchers from developed countries of the West. But in recent years, there has been quite strong demand for the fellowships from many countries, with the exception of the United States, which is still failing to fill its unofficial quota — when the STA scheme was launched in the late 1980s, the Japanese government for political reasons hoped to fill half the places with US scientists, but this goal has never been achieved.

The 5-year plan calls for radical changes and improvements in the government research system in Japan, including a 50 per cent increase in government funding for science and technology and large increases in the number of technical support staff (of which there is a chronic shortage in universities). Thus the research environment in Japan may well be dramatically improved by the end of the five-year period, making it a more attractive destination for non-Japanese scientists.

But opportunities for full-time employment in Japan for non-Japanese are likely to remain limited. Foreign scientists are a tiny minority on the faculty of Japanese universities, and although the private sector was employing fairly large numbers of foreign scientists a few years ago, employment opportunities have declined with the recession of the Japanese economy. Foreign scientists should thus think carefully before making a move to Japan.

David Swinbanks

e-mail addresses/ Web sites

Association Bernard Gregory

abg@grenet.fr

<http://abg.grenet.fr/abg>

ANDES

andes@loire.inapg.inra.fr

hotdocs

<http://enslapp.ens-lyon.fr/~degio/>

hotdocs/index.html